THE UNIVERSITY
of EDINBURGH

Simulation of the Synthesis of Metal-Organic Framework Materials

Naomi Cessford

Doctor of Philosophy

The University of Edinburgh

2013

Abstract

The objective of this work was to develop a molecular simulation method with the capacity to represent the synthesis of metal-organic framework (MOF) structures to the extent of being able to accurately predict the MOF structures that form under specified reaction conditions. MOFs are a class of porous, crystalline solids composed of metal-ion vertices coordinated by organic linker molecules. MOFs are created in a self-assembly process in which the building blocks (reactants) retain their integrity. Under different experimental synthesis conditions, a particular combination of building blocks can react to form differing MOF structures.

The structure of MOFs confers a large degree of tunability, allowing almost limitless potential for the materials to be designed with the capacity to fulfil the requirements of a specific application. Consequently, MOFs have shown promise for a variety of applications including gas storage, separation and catalysis. Thus, the ability to accurately predict the MOF formed by specifying reaction parameters such as temperature, pH and the concentrations of reactants has great potential because, upon identification of a promising hypothetical structure for a particular application, the synthesis conditions ascertained via the simulation method could be used as the basis for the determination of an experimental synthesis procedure. In addition, a simulation method with the capacity to predict MOF structures affords the opportunity to gain a fundamental understanding of the influence of the experimental

synthesis conditions on the structures formed, so as to enable progress towards the rational design of MOFs.

In this work, the experimental synthesis of MOFs via self-assembly is modelled using a kinetic Monte Carlo approach. Ideally, simulation of the self-assembly of the building blocks would be modelled atomistically with all atoms in the reactant and solvent molecules represented explicitly. However, due to the prohibitive computational requirements of such a simulation, in this work a "potential-of-mean-force" (PMF) approach was used to represent the solvent implicitly by encompassing the solvent-mediated behaviour in the interactions between building blocks, thereby reducing the computational cost. The PMF approach to the implicit representation of the solvent involved the utilisation of effective pairwise interactions between the constituents of the reactant species. Following extensive testing to ensure that the explicit-solvent behaviour of the reactants could be replicated using the PMF method, this approach allowed computationally efficient implicit-solvent simulations of the synthesis of MOF materials to be performed.

Thorough assessment of a method developed to simulate the synthesis of MOFs required investigation of a system which, under different reaction conditions, forms differing structures. In this respect, the cobalt succinates represent an unparalleled test because under different reaction temperatures, reactant concentrations, pH and reaction time, seven different phases have been identified. Furthermore, the parameters within which the different phases form have been clearly delineated experimentally. The method developed has been employed, under the appropriate reaction conditions, to simulate the synthesis of two of the seven identified phases of the cobalt succinates. Whilst still subject to computational limitations, the MOF-synthesis simulation method yields structures characteristic of those expected experimentally under corresponding reaction conditions.

Declaration of Originality

I hereby declare that this thesis is of my own composition and that the material contained within has not been submitted for the award of any other degree or professional qualification. The work reported in this thesis is of my own execution, except where due acknowledgement is made in the text.

Naomi Cessford

Acknowledgements

I am grateful to my supervisors, Dr. Tina Düren and Professor Nigel Seaton, for the guidance and support I have received. I am also appreciative of the chemistry-related assistance received from Professor Russell Morris and Professor Paul Wright of the University of St. Andrews. In addition, I am thankful to Dr. Carole Morrison for providing an introduction to the AIM methodology. I gratefully acknowledge the financial support of the Engineering and Physical Sciences Research Council.

Contents

CHAPTER 1

Introduction

The work described in this thesis was undertaken with the broad objective of optimising nanoporous adsorbents for hydrogen purification. This chapter serves to introduce the background information pertinent both to this objective and to the more specific aim of the work, which is subsequently outlined. Additionally, this chapter gives a brief summary of the contents of the following chapters. In establishing the context for the general objective of the work that is subsequently detailed, this thesis commences, in Section 1.1, with a description of the requirement for hydrogen purification.

1.1 Hydrogen Purification

The depletion of global fossil fuel resources, on which over 80 % of the present world energy requirements are based,[1] in conjunction with increasing environmental awareness has led to renewed and growing interest in hydrogen as a clean energy

carrier with a wide range of applications. In addition, there is great demand for hydrogen as a raw material in the production of ammonia, for oil refining, in the manufacture of methanol, and in the chemical and metallurgical industries.[2] However, hydrogen is produced as a by-product in industrial processes including: the production of chlorine and caustic soda in the chlor-alkali process, catalytic reforming in petroleum refineries, and in the manufacture of ethylene.[3] Such by-product hydrogen streams can be recycled, in hydrogen-consuming processes, or, in the absence of better means of utilisation, vented or flared as waste. In view of the development of hydrogen-based technologies, there is compelling incentive to capture the full potential of hydrogen-rich by-product streams by maximising hydrogen recovery and recycling.

The purity of by-product hydrogen is dependent on the origin: whilst hydrogen from the chlor-alkali process is of a high purity, the hydrogen from refineries is only 70 - 80 % pure.[4] Thus, in order to recycle the hydrogen-rich, but impure, by-product process streams, hydrogen purification is required, the degree of which is related to the source of the by-product hydrogen stream and the ultimate application of the purified hydrogen. Hydrogen purification techniques include pressure swing adsorption (PSA), membrane separation and cryogenic separation. To achieve purification of a mixture of gases, PSA involves the separation of the components in the mixture by selective adsorption, whilst membrane processes exploit differences in the selective permeation of different gas molecules, and cryogenic distillation uses differences in the boiling temperatures of components in the mixture to separate impurities.[5] Of the three purification methods, PSA offers greatest flexibility of operation in terms of the ability to maintain hydrogen purity and recovery under varying conditions, and, crucially, yields the highest hydrogen purity: cryogenic and membrane processes typically deliver hydrogen at 90 – 98 volume percent, whereas PSA can produce hydrogen at above 99 volume percent.[6] Thus, PSA is used extensively for hydrogen purification in industrial processes. A comprehensive description of the PSA process for hydrogen purification is provided by Stöcker *et al.*[7] PSA is based on the underlying mechanism of adsorption, the fundamentals of which are described in Section 1.2.

1.2 Adsorption

Adsorption is a process wherein molecules in the gas phase or in solution adhere in a layer of condensed phase to a solid or liquid surface. The molecules that bind to the surface are defined as the adsorbate, whilst the surface on which molecules accumulate is called the adsorbent. The reverse process, in which adsorbate molecules are released from the adsorbent, is termed desorption.[8] A diagram of adsorption and desorption is shown in Figure 1.1.

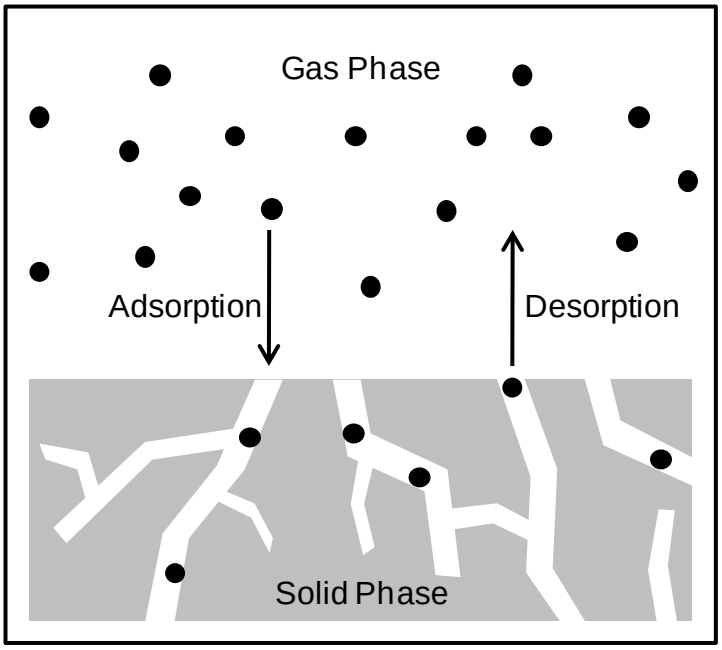

Figure 1.1: Adsorption and desorption processes, where, in this illustration, the adsorbate is gaseous and the adsorbent is a solid material. The adsorbate molecules are represented by filled black circles, the adsorbent is coloured grey and pores in the adsorbent are shown as white regions in the adsorbent.

The adsorption process represented in Figure 1.1 is typical of that involved in PSA, in which gas adsorbate molecules adhere to a solid adsorbent. In the case of adsorption on solid adsorbents, the surface area available for the adherence of adsorbate molecules includes surface irregularities and pore interiors that can be accessed by the adsorbate molecules. The process shown in Figure 1.1 more specifically represents the category of adsorption known as physical adsorption (physisorption), which results from relatively weak interactions (van der Waals

interactions) between the adsorbate and the adsorbent. Due to the weak attraction between the adsorbate molecules and the adsorbent, physisorption is usually readily reversible. A second category of adsorption is chemical adsorption (chemisorption), which involves the sharing or interchange of electrons between the adsorbent material and the adsorbate species. Chemisorbed adsorbate molecules are bonded very strongly to the surface of the adsorbent material and, consequently, are not easily released from the adsorbent. To achieve purification of a mixture of gases, PSA exploits the fact that increasing adsorption capacity is possible with increasing pressure, at a particular temperature, on an adsorbent material. Accordingly, PSA involves steps in which components of the gas mixture are selectively adsorbed and desorbed. Thus, PSA processes generally involve physisorption of adsorbate species.

Adsorption on an adsorbent material results from the adsorbate-adsorbent interactions, which are dependent on the nature of the adsorbate molecule and characteristics of the surface. Therefore, for a gas mixture containing a number of types of adsorbate molecules, there are differences in the affinities of respective components of the mixture for a particular adsorbent. These differences allow the selective adsorption of a particular component in a gas mixture and represent the basis for gas separation or purification by PSA. Thus, the interactions between the adsorbent material and the adsorbate molecules play an important role in the adsorption process. The features of materials typically used as adsorbents are described in Section 1.2.1.

1.2.1 Characteristics of Adsorbent Materials

Since the ability of an adsorbent to discriminate between different adsorbate species dictates the efficacy of the separation of a gas mixture and therefore the gas purification capacity, care is required in the selection of a suitable adsorbent material for a particular separation process. Thus, corresponding to the wide variety of separation and purification applications involving adsorption, there is a vast range of materials that have been successfully employed as adsorbents. Nevertheless, several general attributes are common to the adsorbent materials used in adsorptive

separations. Virtually all solid surfaces can act as adsorbents, with the associated adsorption behaviour dependent on the species in the vicinity. However, because the amount adsorbed is related to the surface area, for effective employment in an adsorption process, an adsorbent material must have a high internal surface that is accessible to the adsorbate molecules.[9] Therefore, the adsorbents used in PSA are porous materials. Porous materials can be categorised, according to the IUPAC notation,[10] in relation to the dimensions of their pores, with materials containing pores with widths less than 2 nm referred to as microporous. Materials containing pores with dimensions between 2 nm and 50 nm are termed mesoporous and those with pores of widths larger than 50 nm are classified as macroporous. In addition to the requirement that adsorbent materials have high internal surface areas and, thus, high adsorption loading capacities, to be suitable for use in PSA separation an adsorbent must be capable of the selective adsorption of one adsorbate species over another. The extent of the ability of a material to separate component A from component B in a mixture is known as the selectivity, $\alpha_{A/B}$, and is defined as:

$$\alpha_{A/B} = \frac{x_A / x_B}{y_A / y_B} \qquad \textbf{1.1}$$

where x_A and x_B are, respectively, the mole fractions of components A and B in the adsorbed phase, and y_A and y_B are the mole fractions of components A and B in the bulk phase, respectively.

Whilst exhibiting different affinities for the species in a mixture so as to enable selective adsorption and therefore separation of a mixture, an adsorbent material must be compatible with the involved adsorbate species to the extent of being chemically inert to all components in the mixture. This compatibility between the adsorbent and the adsorbate species relates to the aforementioned requirement that, in a PSA process, the adsorbate molecules should not react irreversibly with the adsorbent material because the sequential cycles of adsorption and desorption commonly involved in industrial PSA systems necessitate easy regeneration of the

adsorbent material. For effective utilisation, adsorbent materials must also have the capacity to withstand the operating conditions under which the adsorption process is executed.

The outlined characteristics are features of the adsorbent materials utilised in industrial adsorption processes. These adsorbents include amorphous materials such as silica gel and activated carbons, and zeolites, which are crystalline materials. The focus of this work is the class of porous materials known as metal-organic frameworks (MOFs). MOFs are described further in Section 1.3.

1.3 Metal-Organic Frameworks

MOFs have emerged as a class of crystalline porous materials with significant potential for an array of applications including, in addition to gas purification and separation, gas storage and heterogeneous catalysis.[11, 12] Whilst an overview of the structure, synthesis and properties of MOFs is presented in this thesis, comprehensive reviews can be found elsewhere.[13-15] In relation to the aforementioned functions for which MOFs show promise, the specific performance capabilities of a particular MOF are dependent on the intrinsic structure of the material. Details of the structures and related properties of MOF materials are described in Section 1.3.1.

1.3.1 Structure and Properties of MOFs

MOFs are composed of vertices, which are metal ions or metal ion clusters, connected through coordination bonds by organic linker molecules to form regularly ordered networks. Such structures can exhibit uniformly arranged pores and channels, which, generally, following synthesis, are filled with solvent molecules. Frameworks involving strong coordination bonds between vertex and linker species are of sufficient stability that the removal of these guest molecules, by exposing the material to heat or to vacuum conditions, can be tolerated.[16] Therefore, the production of frameworks from linker and vertex ‘building blocks’ can yield porous,

solid structures with large pore volumes and high internal surface areas.[17] The limit of thermal stability of MOFs, however, is lower than that of pure inorganic solids and is typically between 200 and 400 °C in air.[18] That MOFs are thermally less stable than adsorbent materials such as zeolites reflects the weakness of the coordination bonds in MOFs in comparison with the strong covalent bonds in zeolites.[19]

At a fundamental level, the formation of MOF structures from the vertex and linker species can be represented by the schematic shown in Figure 1.2. The topology of the MOF structure produced from a certain combination of vertex and linker species is governed by the coordination preferences directed by the synthesis conditions, and the rigidity of the linker. Because the vertices provide the coordination sites for the organic linkers, the metal coordination geometry, which is related to the oxidation state of the metal ion, determines the geometry of the framework. To bridge the vertices in a MOF structure, the organic linkers have multiple functional groups and are typically di-, tri- or tetravalent ligands. If the ligands are rigid and act as rods between the vertices, robust frameworks with permanent porosity can be produced. The use of ligands with flexible coordination modes has the potential to yield flexible frameworks that can undergo dynamic structural transformations.[20] Thus, in contrast with the synthesis of zeolite-related solids, which involves the use of templates in addition to the solvent and the constituents of the structure, the formation of MOFs does not require the utilisation of specific templating agents.[13]

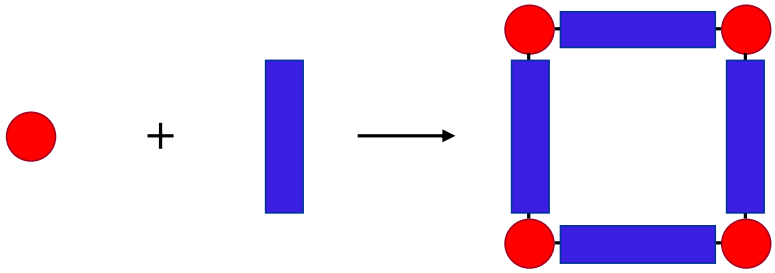

Figure 1.2: Metal-ion vertices and organic linkers combine to form MOFs, where the vertex species is represented by a red circle and the linker is represented by a blue rectangle.

There exist extensive ranges of possible organic linker and metal vertex constituents from which MOFs might be constructed, giving rise to a virtually limitless number of theoretically possible MOF structures.[13] By judicious selection of the linker and vertex species, the modular structural composition of MOFs allows the rational design of the framework structures such that the shapes and aperture sizes of the pores can be tuned, and the pore surface functionalities can be tailored. For instance, the use of differently structured linkers with identical bonding-group arrangements can yield topologically similar MOFs with different pore sizes. This is evident in the IRMOF series of complexes, which consist of ZnO_4-cluster vertices bridged by different linear ditopic carboxylate linkers.[21] In this series, the use of progressively longer linkers enables the production of a number of frameworks based on the same cubic topology but with different pore dimensions and chemical functionalities, and, as a consequence, differing adsorption behaviour.[22]

In addition to the IRMOF structures with cubic networks, a number of the cubic frameworks that incorporate longer linkers have interpenetrated counterparts that are produced under different reaction conditions.[21] Interpenetration is a form of catenation, which is the intergrowth of two or more identical frameworks and is a frequently encountered phenomenon in MOFs. The generation of catenated frameworks can result from the use of long organic linkers but the possibility for catenation is also dependent on the topology of the MOF network.[23] MOF structures are described as interpenetrated when the physically entangled frameworks are maximally displaced from one another. This results in subdivision of the pores and, thus, reduces the pore sizes. Frameworks can also be interwoven, which is another form of catenation, such that there is minimal displacement between the intergrown frameworks.[15]

In addition to the potential to tune the size and shape of the pores in MOF structures, because, in comparison with other porous materials, there is greater correlation between the properties of the constituents of MOFs and the attributes of the resulting structure, there exists great potential for the formation of MOFs with specific pore-surface qualities. Thus, there is the possibility for the construction of MOFs that

exhibit functions such as chirality,[14] magnetism,[24] and luminescence[25] by incorporating building blocks that possess these characteristics. The ability to control the chemical and physical characteristics of the pores in MOFs yields the capacity for the frameworks to be synthesised so as to be selective towards a particular species, as required in adsorptive separation processes. Nevertheless, the ability to produce a MOF structure with a particular functionality is dependent on the synthesis method and the conditions under which the MOF forms. The synthesis of MOFs is detailed in Section 1.3.2.

1.3.2 Synthesis of MOFs

The formation of MOF structures occurs as a single synthetic step because the insolubility of MOFs precludes recrystallisation.[23] In addition, the reaction conditions must provide an environment that facilitates the interaction of the reactant species, and can also satisfy the necessary balance of being mild enough to preserve the functionality of the building blocks yet sufficiently reactive as to allow the formation of the linker-vertex bonds.[15] Therefore, experimentally, MOFs are traditionally generated via "one-pot" synthesis under solvothermal or hydrothermal reaction conditions. In such a MOF-synthesis preparation, the reactants are immersed in either an organic or inorganic solvent contained within an enclosed system under elevated temperature and autogenous pressures.[26] MOFs are usually synthesised at temperatures below 250 °C,[13] and the time taken for formation of a MOF structure from the relevant reacting species ranges from a matter of minutes to several days.[18] In the synthesis of MOFs, the organic ligands are usually introduced directly to the experimental synthesis preparation as reactants or, as is the case for polycarboxylate linkers, input as the corresponding acid. Generally, the source of the metal component is a salt or a metal that undergoes further oxidation in the reaction preparation such that the metal-based vertices form *in situ*.[16]

In addition to the aforementioned solvo- and hydrothermal methods for MOF synthesis, there are several alternative routes to the production of MOFs. Early work relating to the formation of MOFs involved the use of nonsolvothermal synthesis

techniques carried out below the boiling point of the employed solvent in open reaction vessels under ambient pressure.[27] More recently approaches involving microwave-heating, which can allow high-speed synthesis, and sonochemical techniques have been utilised to produce MOF materials.[28] Furthermore, in recent times, high-throughput methods for solvothermal syntheses have been employed to aid in the optimisation of synthesis procedures by examining the influence of separate reaction parameters on the generation of a product material.[29]

The synthesis of MOF materials from the constituent organic linkers and vertices is a self-assembly process in which the building blocks interact to yield the MOF framework.[29] This relates to the fact that, as a result of the vertex-linker interactions, the component building blocks arrange themselves, in solution, to form ordered structures in which the vertices and linkers are connected by highly labile coordination bonds. The lability of the coordination bonds means that the bond formation is rapidly reversible.[30] Thus, coordination bonds between incorrectly connected building blocks can rupture, allowing the reacting entities to rearrange and then reattach to form the thermodynamically favoured structure.[14] This enables the generation of the crystalline architectures inherent to MOFs and prevents the formation of amorphous structures.

Whilst a MOF structure is generated in a single reaction step, to obtain a material suitable for the required application, further processing is generally necessary to remove solvent molecules from the pores in the structure. This constitutes a simple form of post-synthesis modification, the use of which can also enable alteration of the MOF structure following the initial synthesis procedure.[31] Thus, post-synthetic modification of MOFs can be used to introduce certain functionalities to the pores of a structure so that the resulting material has the attributes required for a particular purpose.

As previously stated, the building blocks from which MOFs are synthesised maintain their integrity throughout the synthesis process, such that the characteristics of the constituent species are retained and exhibited in the product framework. In an

experimental preparation, this gives a reasonable degree of control over the nature of the MOF structure produced from the reactants. However, an experimental synthesis procedure is influenced by many properties of the reacting system. For example, in conventional solvothermal syntheses many parameters, such as the starting reactant stoichiometry, the temperature of the system, the solubility of the reactants, the reaction time, and the pH of the solution containing the reactants can have a bearing on the outcome of the synthesis procedure.[13, 29] Thus, under different reaction conditions a given combination of building blocks can form distinct MOF structures. Therefore, for a particular application, whilst the characteristics of a promising hypothetical MOF structure might be identified, experimentally determining the reaction conditions that yield the desired structure is not a straightforward task and can involve a significant degree of trial-and-error. This represents the incentive for the objective of this work, as described in Section 1.4.

1.4 Objective

Motivated by the difficulties associated with ascertaining the experimental synthesis conditions required for the production of a particular hypothetical MOF structure of interest, the objective of this work is to develop a method by which the synthesis of MOFs can be simulated. Given the wide range of applications of MOFs and the ever-increasing demand for new structures, a method that can accurately predict the structure formed when specified building blocks react under particular synthesis conditions affords great advantage. There exist a number of computational approaches to the prediction of crystal structures.[32] The development of these crystal structure prediction approaches stemmed from the desire to guide the design of new materials for which the physical property or characteristic of interest was sensitive to the crystal structure.[33] Crystal structure prediction can, alternatively, assist in evaluating the likelihood, with respect to a molecule with a known crystal structure, of presently undiscovered polymorphism.[34] In general, crystal structure prediction methods are based on the assumption that the crystal structure is determined solely by energy, and that the lowest-energy structure is most likely. Therefore, crystal structure prediction approaches usually constitute efforts in global energy

minimisation, with different methods being distinguishable by the natures of the procedures utilised for the generation of approximate structures and the techniques (most often based on the calculation of lattice energies using density functional theory or force field methods) that are employed as a means of ranking the structures generated.[35]

The most promising computational techniques for crystal structure prediction include simulated annealing, Monte Carlo basin hopping (MCBH), and evolutionary algorithms. Simulated annealing[36] is a powerful stochastic search procedure involving exploration of the crystal structure energy landscape by execution of a random walk (a succession of random steps representing structural alterations) in which there is an ever-decreasing probability of accepting a structural alteration to form a less energetically favourable solution of the crystal structure. The simulated annealing approach has been utilised successfully in determining structures from experimental data, and in predictive studies involving inorganic solids.[35] Monte Carlo basin hopping[37] methods start from a random configuration of the structure and involve successive structural alterations that are accompanied by local minimisation. This transforms the potential energy landscape in to a discrete set of local minimum energy configurations, and, by keeping a history of the sampled minima, can enable faster convergence to the predicted minimum-energy structure. MCBH has been applied successfully in the prediction of the crystal structures of clusters up to 100 atoms.[38] Energy-landscape explorations via simulated annealing and MCBH methods start from a single point, resulting in the possibility that not all of the low-energy configurations are sampled, and necessitating the execution of multiple runs from different starting configurations. This issue is not a problem associated with evolutionary algorithm methods, which work on a set of samples.[35] Evolutionary algorithm[39] methods involve moves (representing structural alterations) that imitate aspects of biological evolution such that the starting set is gradually improved (in terms of the energy of the configuration) by performing successive operations analogous to the evolutionary processes of crossover (which involves combining structural motifs in different candidate structures in the set sampled to form a new structure), mutation (which involves performing random displacements of atoms in a

sampled structure to generate a new structure), and natural selection (which involves removing less favourable configurations from the set sampled). Thus, sampled structures displaying favourable features prosper and those with less favourable features diminish, with the result that the most energetically favourable structure should evolve. There exist many variations of the exact natures of the operations central to evolutionary algorithms and the capabilities of these methods have been demonstrated in a variety of situations, including in the structure prediction of metallic and alloy clusters,[40] and in the solution of both dense and microporous oxide structures.[41]

The described methods represent a small proportion of the many crystal structure prediction approaches that have been formulated. In recent years progress in the field of crystal structure prediction has been evaluated by carrying out blind prediction tests.[34, 42-45] These tests involve the assessment of predictions, made by several participating groups employing a range of computational methodologies, of the structures of a number of organic compounds for which the experimental crystal structures have been accurately determined. As a means of judging the progress of crystal structure prediction methods, the periodic execution of blind tests is, however, limited by the fact that the results of the tests can be influenced by the natures of the molecules selected for structure prediction in addition to the methodological advances made between successive tests.[44] Nevertheless, the results of the blind tests have shown that significant advances have been made in recent times. The need for further work in the field of crystal structure prediction is clear, however.

In relation to MOFs, the AASBU (automated assembly of secondary building units) method, which involves simulated annealing, has previously been shown to be capable of structure solution.[46] This approach involves the combination of secondary building units, using physically reasonable rules, to compile a library of plausible structures. In the work described in this thesis, the aim is to represent the synthesis of MOFs by developing a simulation method that follows a reaction pathway that is

dependent on the relevant synthesis conditions, and, thus, obtain a means of predicting the MOF structure that forms under certain conditions.

In pursuit of this aim, work was undertaken in several distinct areas. The work carried out in relation to each topic is described separately in the following chapters, the contents of which are summarised in Section 1.5.

1.5 Thesis Outline

Following the introduction to this thesis and the definition of the objective of this work given in this chapter, Chapter 2 (Simulation of the self-assembly of metal-organic framework materials) contains details of the simulation methodologies employed in this work and presents the MOF system that is focussed on as a means of testing a method developed to simulate the synthesis of MOFs. Further, in Chapter 2, details are given of an initial simulation of the self-assembly of the components of the test system.

Chapter 3 (Calculation of the potential of mean force between ionic species) presents the development of an alternative means of representing the solvent, for application in simulations of the synthesis of MOFs. Thus, Chapter 3 contains details of both the underlying theory behind this solvent-representation method and the procedures involved in the development of the method.

The capacity of the method developed for representing the solvent in simulations of the synthesis of MOFs was assessed in Chapter 4 (Evaluation of the potential-of-mean-force approach). This was achieved by comparing the behaviour of the building blocks involved in the test system in simulations involving, respectively, the initial solvent representation, and the developed solvent representation.

Chapter 5 (Implicit-solvent simulation of the synthesis of MOF materials) contains a description of the simulation method developed to represent the synthesis of MOFs, and, further, gives details of the implementation of the method. With application of

the simulation method to the test system utilised in this work, this chapter also outlines the ability of the method to predict the MOF structures that form under specified reaction conditions.

Whilst Chapter 6 (Evaluation of the applicability of ideal adsorbed solution theory to mixture adsorption in MOFs) is not specifically related to the primary objective of this work, this chapter contains a thorough evaluation of the capacity of a means of predicting mixture adsorption in MOFs. Additionally, this chapter serves to highlight the impact of characteristics of the structures of MOFs on the adsorption of different mixtures of gases.

The overall findings and outcomes of the contents of Chapters 1-6 of this thesis are discussed in Chapter 7 (Conclusions and further work). This chapter also contains recommendations for future work.

The results of both work described herein and undertakings executed in association with the work detailed in subsequent chapters have been published and/or presented in a number of sources. These sources are specified in Section 1.6.

1.6 Publications and Presentations

Publications

"Evaluation of ideal adsorbed solution theory as a tool for the design of metal-organic framework materials", N. F. Cessford, T. Düren and N. A. Seaton, *Industrial and Engineering Chemistry Research*, 2012, 51, 4911-4921

"The synthesis, structures and reactions of zinc and cobalt metal-organic frameworks incorporating an alkyne-based dicarboxylate linker", A. D. Burrows, L. C. Fisher, D. Hodgson, M. F. Mahon, N. F. Cessford, T. Düren, C. Richardson and S. P. Rigby, 2012, *CrystEngComm*, 2012, 14, 188-192

"Protecting group and switchable pore-discriminating adsorption properties of a hydrophilic-hydrophobic metal-organic framework", M. I. H. Mohideen, B. Xiao, P. S. Wheatley, A. C. McKinlay, Y. Li, A. M. Z. Slawin, D. W. Aldous, N. F. Cessford, T. Düren, X. B. Zhao, R. Gill, K. M. Thomas, J. M. Griffin, S. E. Ashbrook and R. E. Morris, *Nature Chemistry*, 2011, 3, 304-311

Presentations

N. F. Cessford, T. Düren and N. A. Seaton
"Evaluation of the applicability of the ideal adsorbed solution theory to mixture adsorption in metal-organic framework materials" poster presentation
BZA Annual Meeting, August 2009, Ambleside, UK

N. F. Cessford, T. Düren and N. A. Seaton
"Building blocks for hydrogen purification: simulation of the synthesis of metal-organic framework materials" poster presentation
School of Engineering Postgraduate Conference, May 2010, University of Edinburgh, UK

CHAPTER 2

Simulation of the self-assembly of metal-organic framework materials

Self-assembly is the fundamental mechanism that controls the formation of well-defined structures from disordered parts. With minimisation of the free energy of the system as the driving force, self-assembly can, through spontaneous and reversible organisation of components, yield ordered structures which represent the thermodynamic minimum, or an approach to the thermodynamic minimum configuration.[47] Specifically, when the constituents involved in self-assembly are molecules, the process is labelled molecular self-assembly. Self-assembly in a molecular system is defined by a number of distinguishing characteristics:[48] firstly, the self-assembled system must have a higher degree of order than the isolated components; secondly, the interactions between the components that self-assemble are generally weak and non-covalent, commonly taking the form of van der Waals and Coulombic interactions, hydrophobic interactions and hydrogen bonding, but

weak covalent bonds (coordination bonds) have also been acknowledged as being valid in self-assembling systems;[49] thirdly, for self-assembly to yield ordered structures, the components either must associate reversibly or must be allowed to alter their positions within the self-assembled entities formed. The opportunity for the constituents of the self-assembled entities to undergo repositioning or to reversibly associate is necessary for realisation of the thermodynamically favoured structure. In addition, to achieve the mobility required for the components to assemble to form ordered structures, self-assembly generally occurs in solution or at an interface. The environment in which the components interact, and self-assembly therefore occurs, has a significant bearing on the process. Thus, the production of regular, ordered MOF structures via the formation of coordination bonds between metal vertices and organic linkers that interact in a solvent clearly represents a self-assembly process.

In order to develop a simulation method to accurately predict the MOF structure resulting from a particular set of reaction conditions, the self-assembly process must be suitably represented in the simulation. Computer simulation represents a tool by which molecular systems can be studied so as to explore macroscopic properties by evaluating the behaviour at a microscopic level.[50] A computer simulation is the execution of a model which embodies the characteristics of the system under investigation, where execution involves the calculation of numerical values for specific properties at certain thermodynamic states and is effectively a 'computer experiment'. The macroscopic properties are connected to the microscopic simulation via statistical mechanics, which represents a bridge between the microscopic variables that characterise the individual molecules of the system, and the macroscopic behaviour of the entire system.[51, 52] The principles of statistical mechanics are expounded in great detail in volumes by McQuarrie[51] and Hill.[53] In statistical mechanics, the properties of a system are determined by constructing a large number of replicas of the system, each representative of a particular microscopic state. Together the replicas constitute an ensemble, i.e. a collection of configurations of which each represents a point in the phase space, which is defined as the space containing all of the possible values of the position and momentum of

each molecule in the system. The selection of an appropriate ensemble is dependent on the conditions of the physical process under investigation. In relation to systems which self-assemble, the macroscopic variables that remain constant are the number of molecules involved, *N*, the volume, *V*, in which the system undergoes self-assembly, and the temperature, *T*, at which the process occurs. Therefore, the natural choice of method for the simulation of self-assembly is the canonical or NVT ensemble. The canonical ensemble is introduced in Section 2.1.

2.1 Canonical Ensemble

Characterised by a fixed total number of molecules, a given volume, and a specified temperature, the canonical ensemble[54] relates to the situation in which the system is not isolated from the surrounding environment, but rather is maintained in thermal equilibrium with a much larger system, known as a reservoir, which imposes on the system the specified temperature. A schematic representation of a system in a canonical ensemble is presented in Figure 2.1.

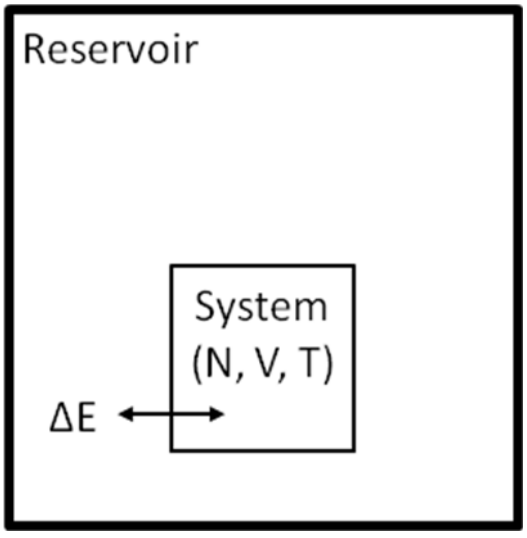

Figure 2.1: The system in contact with an energy reservoir. The bold border line enclosing the reservoir represents the thick boundary isolating the composite system from the environment, and the faint border line surrounding the system represents the thin, rigid, diathermal, impermeable boundary through which energy is exchanged.

In the context of thermodynamics and statistical mechanics, the reservoir constitutes a body that, compared to the system under investigation, is sufficiently large so as not to experience an appreciable change in temperature as a result of exchanging heat with the system. The walls of the system are rigid, such that the volume is fixed, and impermeable to molecules, so that the number of molecules remains constant. The energy of the system is not fixed by the contact with the reservoir and can, therefore, vary over all the possible energy values associated with the system. Thus, in the canonical ensemble, each microstate is associated with a particular energy value.

The probability, P_i, of finding the system in a particular microstate i with energy, E_i, is given by:

$$P_i = \frac{1}{Q_{NVT}} \exp(-\beta E_i) \qquad \textbf{2.1}$$

where $\beta = 1/k_B T$, in which identity k_B represents the Boltzmann constant. Q_{NVT} is a normalisation constant defined as the canonical partition function. The exponential term in Equation 2.1 is known as the Boltzmann factor and represents the weighting attached to a given microstate. Thus, Equation 2.1 shows both that the probability with which a particular microstate is visited depends on the energy of the microstate and the temperature of the system, and that the probability associated with a particular microstate is not related to the properties of the energy reservoir beyond the involvement of the reservoir in the establishment of the temperature of the system. In addition, that the probability of occurrence of a specific microstate is proportional to an exponential function of the energy indicates that high-energy microstates will be visited exponentially less frequently than microstates with low energies. The canonical partition function is given by the summation of the Boltzmann factors of all accessible microstates of the system, as shown in Equation 2.2.

$$Q_{NVT} = \sum_i \exp(-\beta E_i) \qquad \textbf{2.2}$$

The canonical partition function serves as a bridge between the canonical ensemble and the key thermodynamic properties of the system. Since the canonical partition function and the Helmholtz free energy, A_{NVT}, are both associated with the natural independent variables of temperature, volume and number of molecules, the relation shown in Equation 2.3 can be derived.

$$A_{NVT} = -k_B T \ln Q_{NVT} \qquad \textbf{2.3}$$

The Helmholtz free energy is a fundamental property of the system, from which all other thermodynamic properties can be determined. Thus, the partition function represents a means by which the macroscopic properties of the system can be evaluated from the characteristics of the system at the microscopic scale. The microscopic behaviour of the system can be studied via simulations involving the generation of microstates, so as to construct an ensemble.

2.2 Simulation Techniques

The simulation techniques by which the microstates can be produced are now considered. Computer simulations in statistical mechanics are based on two types of techniques: those based on deterministic approaches, molecular dynamics (MD), and those founded on stochastic methods, Monte Carlo (MC). In addition, there are a number of different hybrid methods which incorporate features of both MD and MC techniques.[55] MC and MD simulation techniques involve the generation of a sequence of microstates representing the appropriate statistical ensemble for the relevant variables. Both methods represent cases of importance sampling, such that the computational effort is concentrated on creating microstates that are representative of the equilibrium ensemble, rather than merely sampling the entire phase space.[56] This effects a vast improvement in efficiency, which makes computer simulations involving statistical mechanical models practicable. The principles directing the way in which the system moves from one microstate to another are described for MD and MC simulations, respectively, in Sections 2.2.1 and 2.2.2.

2.2.1 Molecular Dynamics

MD is described in detail by Allen and Tildesley.[57] In MD simulations, microstates are generated through the solution of Newton's equations of motion for all of the atoms in the system. The relevant equation of motion for atom *j*, with mass, m_j, and position vector, $\mathbf{r}_j$, is expressed as follows:

$$\mathbf{F}_j = m_j \frac{d^2\mathbf{r}_j}{dt^2} \qquad \textbf{2.4}$$

where $\mathbf{F}_j$ is the force vector for atom *j* and can be written as the negative of the gradient of the potential energy function, $U(r_1,...,r_N)$, through which the *N*-body systems interacts, as shown in Equation 2.5.

$$\mathbf{F}_j = -\nabla_j U(r_1,...,r_N) \qquad \textbf{2.5}$$

The differential equation produced by combining Equations 2.4 and 2.5 can be solved, yielding a trajectory describing the variations with time of the positions, velocities and accelerations of each atom. Thus, the temporal evolution of the system can be tracked from the initial positions, orientations and momenta assigned to the molecules by successively applying the forces for a small time step and updating the atomic configurations and parameters using an integration algorithm. The time step in MD simulations is typically of the order of femtoseconds (10^{-15} s), as dictated by the fastest atomic motions. Such a time step limits the time scale accessible by MD simulations. In MD, the molecular motions constituting the evolution of the system occur naturally under the influence of the specified potential energy function. However, the accuracy of the molecular motions is dependent on the selection of the potential function.

2.2.2 Monte Carlo

MC simulations are described in detail by Frenkel and Smit.[56] MC simulations involve the generation of a series of microstates by making successive random trials, altering the molecular configuration of the system following a Markov chain approach in which a new microstate i_{n+1} is generated stochastically from the preceding microstate, *i*, so that no information relating to the previous configurations i_{n-1}, i_{n-2}, ... is required. MC simulations typically incorporate the importance sampling method introduced by Metropolis *et al*.,[58] such that a new configuration generated by a random change to the previous configuration is accepted or rejected according to a criterion devised to produce the desired probability distribution. For the canonical ensemble, which, as previously described, is suited to the simulation of self-assembly, the trial moves attempted involve the translation of molecules and, for non-spherical species, the rotation of molecules. The probability of acceptance, *p*, is based on the potential energy change, $\Delta U = U(n) - U(o)$, associated with the random move from old, *o*, to new, *n*, configurations, such that

$$p = \min\left[1, \exp\{-\beta(\Delta U)\}\right] \qquad \textbf{2.6}$$

Thus, if the energy change associated with the random movement of a molecule is negative, the move is always accepted and is added to the ensemble. A positive energy change results in the move being accepted with probability $\exp\{-\beta(\Delta U)\}$. Rejection of a trial move results in the molecule concerned being retained at the old position, and the old configuration being added to the ensemble.

Rather than evolving smoothly in correspondence with the equations of motion, as per MD simulations, MC simulations sample phase space by moving abruptly from configuration to configuration, and so cannot yield the evolution of the system with respect to time. In addition, since MC simulations need not follow a realistic path between sampled configurations to obtain the correct equilibrium probability distribution, there is greater flexibility with respect to the choice of moves that can be implemented to most effectively sample phase space.[59] Therefore, in MC

simulations, the stochastic transitions between microstates might or might not be physically realisable as long as detailed balance, also known as microscopic reversibility, is obeyed. Adherence to the detailed balance criterion demonstrates that the sampling in the MC simulation is uniform and unbiased, and is proof that the correct distribution is sampled. When detailed balance is enforced, the Markov process is guaranteed to be reversible, a feature that is essential for accurate representation of the self-assembly process. Detailed balance implies that, at equilibrium, the probability of being in a given state *o* and transitioning to state *n* is equivalent to the probability of being in state *n* and moving to state *o*. The detailed balance criterion is expressed as:

$$N(o)\pi(o \rightarrow n) = N(n)\pi(n \rightarrow o) \qquad \textbf{2.7}$$

where *o* represents the old state, *n* signifies the new state, and π symbolises the transition between the old and new states. Thus, as shown in Equation 2.7, the rejected moves are implicated in the attainment of detailed balance.

In comparison with MD simulations, a greater variety of ensembles can be simulated using MC, with the trial moves attempted depending on the ensemble of interest. In order to determine whether to accept or reject the configurations resulting from the trial moves in MC simulations all relevant interactions must be evaluated. Therefore, as with MD simulations, the use of an accurate potential function is critical to the success of the simulation. Due to their inherent randomness, MC simulations are naturally ergodic, such that the entire phase space is sampled during the progress of the simulation. In contrast, the deterministic nature of MD simulations means that accomplishing ergodicity is a more challenging task.[59]

2.3 Simulation Methodology for MOF Self-Assembly

Since the primary objective of this work is to develop a method that can predict the MOF structures resulting from specified reaction conditions, the time dependence of

the process is of little concern. Therefore, the MC method was selected as a means of conducting the self-assembly simulations. In addition to not being subject to time limitations, the MC approach is more flexible with regard to the implementation of new move types, so that, as is necessary in order to simulate the synthesis of MOFs, a move representing the reaction between the reactants from which the MOF is generated can be more easily developed.

In addition to statistical theory, MC simulations involve a number of techniques that are employed to ensure successful application, and to save computational effort. The general objective of MC simulations is to yield information about molecular systems at the macroscopic level. However, in any simulation, the number of molecules is limited to that for which the desired properties can be obtained in a reasonable computation time. Thus, the number of molecules involved in a MC simulation is far removed from the number typical of a thermodynamic system, which is of the order of Avogadro's constant (10^{23}). At an equivalent density to the macroscopic system of interest, the number of molecules involved in the simulation will be contained within a small dimension of the system, with the effect that a significant proportion of the molecules in the system will be influenced by the boundaries of the system.[60] A simulation of a system in which a large proportion of the molecules are influenced by the boundaries would not be characteristic of a small volume constituent of the thermodynamic system. Within the MC simulation approach, the impact of the system boundaries can be dealt with, as described in Section 2.3.1, by imposing periodic boundary conditions (PBCs) on all faces of the simulation cell.

2.3.1 Periodic Boundary Conditions

The application of periodic boundary conditions allows a finite system to approximate the properties of an infinite system, by eliminating the finite-size effects resulting from the influence of the boundaries.[61] PBCs solve the finite-size problem by generating image copies of the original simulation cell, which periodically repeat the system in every direction, so as to preserve macroscopic homogeneity and

effectively represent an infinite system. The copies are virtual representations of the unit cell, and the molecules included in the copies are not explicitly modelled.

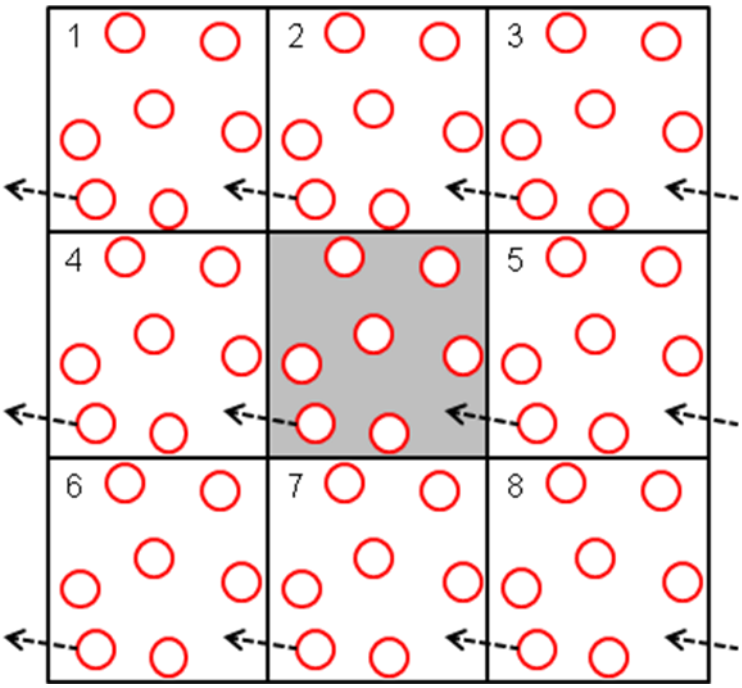

Figure 2.2: A two-dimensional representation of a system in which periodic boundary conditions are applied. The original simulation cell is shaded.

As shown in Figure 2.2, when simulations are carried out using PBCs, when a molecule is moved such that it exits one side of the central simulation box, a corresponding image molecule re-enters the opposite face of the central box. Thus, the application of PBCs ensures both that the simulation cell always contains the correct number of molecules and that the molecules are free to move without being influenced by the boundaries of the system. For the purposes of calculating the potential energy of a particular microstate sampled in a MC simulation, PBCs ensure that each molecule in the system is surrounded by the correct neighbouring molecules, which might belong either to the original unit cell or to the periodic images of the original simulation box. However, for accurate evaluation of the potential energy of a given configuration the correct interactions must be considered. In a periodic system, the minimum image convention can be employed to guarantee the inclusion of the appropriate interactions in the calculation of the potential energy of a given microstate, as explained in Section 2.3.2.

2.3.2 Minimum Image Convention

A consequence of the application of PBCs in a simulation is that a molecule will "see" multiple copies of itself and of each of the surrounding molecules. However, the interaction between a particular molecule and a certain neighbouring molecule should only be counted once. In simulations, in which interactions between molecules are typically calculated as the sum of the pairwise interactions between the atom constituents of the molecules, the minimum image convention is implemented such that, for a particular atom, only interactions with the closest images (the minimum image) of each of the neighbouring atoms will be considered, regardless of whether the surrounding atoms exist in the original simulation box or in one of the periodic image boxes.[62] This can be achieved by ensuring that atoms only interact with neighbouring atoms closer than half of the shortest dimension of the simulation cell. Thus, for simulations in which the minimum image convention is applied, the effects of the artificial periodicity imposed on the system by PBCs will be indiscernible.

Thus, in simulations, the application of PBCs and the minimum image convention represents a suitable methodology for the calculation of the interactions involved in computations of the potential energy. However, the form of the interactions involved in the calculation is also crucial to the accuracy of the evaluation of the potential energy values required in simulations. Therefore, the utilisation of appropriate molecular potentials, through which the interactions are described, is a key feature in the realistic simulation of the self-assembly process. The details of the molecular potentials employed in this work are outlined in Section 2.4.

2.4 Molecular Potentials

As described in Section 2.2.2, accurate calculation of the potential energy associated with the sampled configurations is fundamental to the success of MC simulations. In the MC approach, the molecular potential energy reflects how a molecule interacts both with itself and with the other molecules in the system, and is generally

represented as the sum of the contributions resulting from intermolecular (non-bonded) interactions and the intramolecular (bonded) interactions. The relevant interactions are modelled using molecular potentials, in which the energy is specified by means of an analytical expression relating the energy to the geometry of the atomic configuration. The functional forms of the molecular potentials are selected so as to emulate the forms of the corresponding real interactions. Constants in the functions are parameterised by fitting certain properties in test simulations to the equivalent properties obtained from either experimental sources or high-level quantum mechanical calculations.[63]

Molecules are generally represented in simulations as a collection of interaction points, which frequently relate to the positions of the atoms in the molecule, but might also correspond to a definable position in the molecule if so required for the accurate representation of the features of the molecule. Molecules can also be modelled such that a group of the constituent atoms are represented by a single interaction site, known as a "united atom". Such entities can effect a substantial reduction in the number of interactions that must be evaluated in the simulation. Thus, evaluation of the potential energy of a molecule involves consideration of both bonded and non-bonded interactions for all relevant interaction sites. The details of the intermolecular and intramolecular interactions required for evaluation of the molecular potential energy are described, respectively, in Sections 2.4.1 and 2.4.2.

2.4.1 Intermolecular Interactions

Intermolecular, or non-bonded, interactions describe the alteration in the potential energy resulting from changes in the external position of a molecule. Such interactions are usually associated with van der Waals and electrostatic energy contributions. Typically, molecular potentials specify such interactions as occurring between pairs of interaction sites, with the magnitude of the potential related to the separation distance between two individual sites. Thus, the potential energy of a microstate sampled in a MC simulation is determined by assuming pairwise additivity, such that the total energy of a collection of species is calculated as the

sum of the interactions between pairs of species considered to be isolated from all other molecules.[64] The representations of van der Waals interactions and electrostatic interactions are described, respectively, in Sections 2.4.1.1 and 2.4.1.2.

2.4.1.1 van der Waals Interactions

The van der Waals interactions between interaction sites are represented by means of the Lennard-Jones (LJ) 12-6 potential, which encompasses both long-range dispersion interactions and the short-range repulsive interactions that result from the overlap of electron clouds. The LJ potential, U_{LJ}, as a function of the distance, r_{ij}, between interaction sites i and j has the form:

$$U_{LJ}(r_{ij}) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \qquad \textbf{2.8}$$

where ε_{ij} and σ_{ij} are known as the LJ parameters and represent, respectively, the depth of the LJ potential well and the collision distance for interaction sites i and j. The functional form of the LJ potential is shown in Figure 2.3.

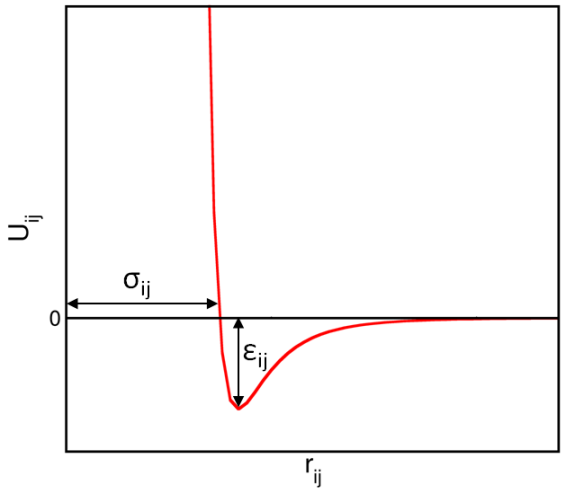

Figure 2.3: The functional form of the Lennard-Jones 12-6 pair potential.

In the specification of a molecule, the parameters associated with the individual interaction sites contained within the molecule are detailed. However, for the

interaction between dissimilar interaction site types, the LJ parameters are generally calculated according to a set of rules, such as the Lorentz-Berthelot mixing rules, shown in Equations 2.9 and 2.10 for calculation of the cross-site sigma and epsilon values, respectively.

$$\sigma_{AB} = \frac{1}{2}(\sigma_{AA} + \sigma_{BB}) \quad \textbf{2.9}$$

$$\varepsilon_{AB} = \sqrt{\varepsilon_{AA}\varepsilon_{BB}} \quad \textbf{2.10}$$

Thus, for the interaction between sites *A* and *B*, the collision diameter, σ_{AB}, is equal to arithmetic mean of the values for the individual sites, and the well-depth ε_{AB} is represented by the geometric mean of the values for the respective interaction sites.

2.4.1.2 Electrostatic Interactions

The electrostatic interactions are modelled using the Coulombic potential:

$$U_{coul}(r_{ij}) = \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \quad \textbf{2.11}$$

where q_i and q_j are the magnitudes of charges associated with interaction sites *i* and *j* separated by r_{ij}, and ε_0 is the vacuum permittivity. The partial charges involved in the calculation of electrostatic interactions are assigned to the interaction sites in a molecule so as to approximate the molecular charge distribution. Where appropriate, molecular potentials can specify the placement of the partial charges such that the electric moments of the molecule (eg. dipole, quadrupole) can be reproduced.[65]

Electrostatic interactions decay slowly with order r^{-1} and are, therefore, long-ranged compared to the van der Waals interactions, which decay with order r^{-6}. Thus, whilst the short-ranged van der Waals interactions can be effectively calculated using a spherical cut-off method, this treatment is liable to introduce errors if applied to the evaluation of the longer-ranged electrostatic interactions.[66] The application of a

simple spherical cut-off scheme to an interaction site involves the specification of a distance that corresponds to the limit to the radial distance within which interaction sites must exist to be included in the pairwise summation calculation of the potential. With respect to electrostatic interactions, the application of a spherical truncation can produce a discontinuity near the cut-off distance.[67] Consequently, a number of alternative methods have been developed for handling electrostatic interactions. The most basic of these are schemes involving the application of shifting or switching functions, which modify the potential so as to attempt to minimise the effect of discontinuities that result from abrupt truncation. The purpose of the implementation of such functions is to alter the coulombic potential so that the potential is reduced to zero at the cut-off distance.[68] A more efficient approach to the treatment of electrostatic interactions in systems with PBCs is the Ewald summation technique. Ewald summation was developed by Ewald[69] in 1921, and is described in detail elsewhere.[56, 70] Ewald summation involves the division of the slowly convergent Coulombic energy sum into two rapidly convergent components. The first term is short-ranged and handled in real space, whilst the second term is long-ranged and dealt with in reciprocal space. The splitting of the Coulombic sum is accomplished by means of shielding charge distributions centred at the same locations in space as the charges associated with interaction sites, but with opposite charges. The shielding charge distributions give rise to the short-range component because, by shielding the charges, the electrostatic interactions decay rapidly and can be evaluated through the application of a spherical cut-off. The shielding distributions are cancelled by adding a second charge distribution with equal magnitude but opposite sign to those in real space. This represents the second, long-ranged component of the sum, which is a smooth function with a rapidly decaying Fourier transform. The combination of the components in real space and reciprocal space yields the original charge distribution. The charge distributions in the original, real and reciprocal spaces are illustrated in Figure 2.4.

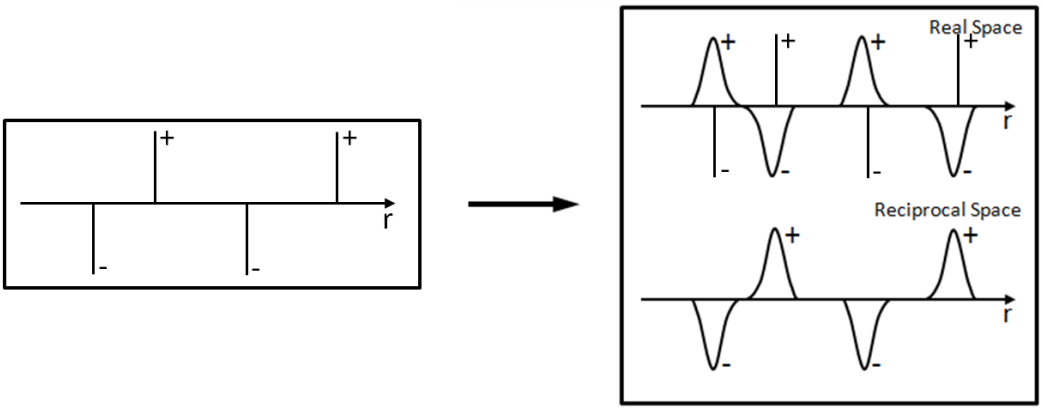

Figure 2.4: Charge distributions in Ewald summation.

The Ewald summation technique has been implemented for evaluation of electrostatic interactions in a range of systems, including condensed systems.[71] However, the application of the Ewald method is limited by the considerable computational cost associated with the calculation in reciprocal space.[68] Several variations of the standard Ewald method have been successfully implemented, including the particle mesh Ewald (PME)[72] and smooth particle mesh Ewald (SPME).[73]

A further technique for the evaluation of electrostatic interactions is the pairwise summation method proposed by Wolf *et al.*[74] The central concept on which the Wolf method is based is that the poor convergence of the Coulombic energy summation results from the non-neutrality of the local charge within a spherically truncated system, with radius r_{cut}. Thus, the method involves the addition of a correcting term, which guarantees that the net charge within a cut-off sphere is equal to zero, to the direct pairwise summation of the spherically truncated Coulombic potential for a system containing N charges. As shown in Equation 2.12, the energy contained within the charge-neutralised sphere, $U_{charge\ neutral}$, can be calculated by subtracting the unbalanced sphere-surface charges from the traditional Coulombic sum.

$$U_{charge\ neutral}(r_{ij}) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq i\ r_{ij}<r_{cut}}\frac{q_i q_j}{r_{ij}} - \lim_{r_{ij}\to r_{cut}}\frac{q_i q_j}{r_{ij}} \qquad \textbf{2.12}$$

Computation of the charge-neutralised energy in a sphere, as demonstrated by Adams,[75] approximates shifting the correct energy by a constant, which represents the self-interaction term, as given in Equation 2.13.

$$U_{shift}(r_{ij}) = \frac{1}{2r_{cut}} \sum_{i=1}^{N} q_i^2 \qquad \textbf{2.13}$$

Thus, for the spherically truncated system, the approximate correct energy, U_{Wolf}, is given by subtracting Equation 2.13 from Equation 2.12.

$$U_{Wolf}(r_{ij}) = \frac{1}{2} \sum_{i=1}^{N} \sum_{j \neq i \; r_{ij} < r_{cut}} \frac{q_i q_j}{r_{ij}} - \lim_{r_{ij} \to r_{cut}} \frac{q_i q_j}{r_{ij}} - \frac{1}{2r_{cut}} \sum_{i=1}^{N} q_i^2 \qquad \textbf{2.14}$$

Equation 2.14 converges to the correct values of the Coulombic sum for high values of r_{cut}, and could, in theory, be used to evaluate the electrostatic interactions in a simulation. However, in practice, the cut-off value required would be of such magnitude as to be impracticable.[76] Therefore, to make the Wolf method computationally efficient, a damping function, which takes the form of the fast-decaying complementary error function, erfc, is applied, yielding the summation equation for the Wolf potential, U_{Wolf}, given in Equation 2.15:

$$U_{Wolf}(r_{ij}) = \frac{1}{2} \sum_{i=1}^{N} \sum_{j \neq i \; r_{ij} < r_{cut}} \frac{q_i q_j}{r_{ij}} (\mathrm{erfc}(\alpha r_{ij})) - \lim_{r_{ij} \to r_{cut}} \frac{q_i q_j}{r_{ij}} (\mathrm{erfc}(\alpha r_{ij})) - \left(\frac{(\mathrm{erfc}(\alpha r_{ij}))}{2r_{cut}} + \frac{\alpha}{\sqrt{\pi}} \sum_{i=1}^{N} q_i^2 \right) \qquad \textbf{2.15}$$

where α is a damping parameter. In addition to being easier to implement, the Wolf method can significantly reduce computational time in comparison with Ewald summation, without sacrificing accuracy.[77] Thus, there exist a number of different methods by which electrostatic interactions can be evaluated in MC simulations.

Whilst the intermolecular interactions describe the potential between molecules in a simulation, the interactions between atoms belonging to the same molecule are also important for the accurate representation of a molecular species. The methods by which intramolecular interactions are represented are discussed in Section 2.4.2.

2.4.2 Intramolecular Interactions

Intramolecular, or bonded, interactions describe the changes in the molecular potential energy resulting from alterations in the internal configuration of a molecule. Within a molecule, the intramolecular interactions are typically calculated between interaction sites that correspond to the positions of the atoms. The internal structure of a molecule is described by the bond lengths between two atoms, the bond angles characterised by the positions of three atoms, and the dihedral angles defined by three consecutive bonds between atoms. Therefore, bonded interactions include contributions arising from bond stretching, bond angle bending and torsional rotations. Representations of the intramolecular interactions described in molecular potentials are illustrated in Figure 2.5.

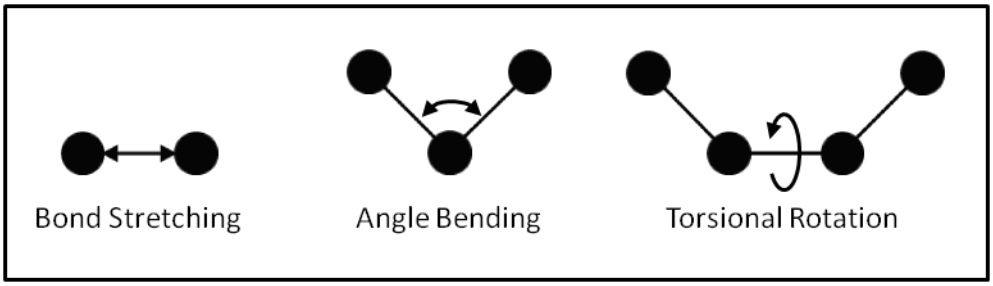

Figure 2.5: Types of intramolecular interactions.

The stretching of the bond length between two atoms in a molecule is modelled using a harmonic potential:

$$U_{bond}(r_{ij}) = \frac{1}{2} k_b (r_{ij} - r_b)^2 \qquad \textbf{2.16}$$

where k_b is the force constant associated with the stiffness of the bond, r_b is the equilibrium bond length between the two atoms in the structure, and r_{ij} is the actual distance between the bonded atoms. The harmonic form of the potential means that an increase in the energy associated with the bond, U_{bond}, will result from distortion from the equilibrium bond length. The function describing how the bond angle energy, U_{angle}, varies with deformation away from the equilibrium position is also harmonic in nature, as shown in Equation 2.17:

$$U_{angle}(\theta) = \frac{1}{2} k_\theta (\theta - \theta_0)^2 \qquad \textbf{2.17}$$

where k_θ is the force constant associated with the angle, θ_0 is the equilibrium value of the bond angle, and θ is the actual bond angle. The third type of intramolecular interaction describes the variation in energy of a molecule as a result of torsional rotation about one of the bonds in the molecule. The torsional angle potential energy, $U_{torsional}$, can be expressed as shown in Equation 2.18:

$$U_{torsional}(\phi) = \frac{1}{2} k_\phi [1 + \cos(n\phi - \delta)] \qquad \textbf{2.18}$$

where ϕ is the torsional, or dihedral, angle, k_ϕ is the torsional force constant, n is the periodicity of the angle, and δ is the phase of the angle. In addition to the potentials associated with the internal bond lengths, bond angles and dihedral angles in molecules, in some molecular potentials, additional van der Waals and electrostatic interactions are included between pairs of atoms belonging to the same molecule. Such interactions are calculated between atoms separated by three or more consecutive bonds and are, thus, known as 1-4 intra-pair intramolecular interactions.

Thus, the details of both the methods of representing molecules and the techniques involved in simulations have been described. In the development of a method to simulate the synthesis of MOFs, evaluation of the predictive capacity of the method necessitates assessment for a system for which more than one MOF structure can be

synthesised through the application of different reaction conditions. The system for which the method was tested is described in Section 2.5.

2.5 MOF System

To thoroughly assess, and have confidence in, the validity of a method developed to simulate the synthesis of MOFs, the cobalt succinates are used as a test system. The cobalt succinates represent an unparalleled class of MOF materials by virtue of the fact that, under different reaction temperatures, reactant proportions, pH and reaction time, seven different phases have been identified. In addition, for the cobalt succinate system, the parameters within which each of the seven phases form have been clearly delineated experimentally, thereby providing the capacity to effect a precise test of the method developed.

The capacity for multiple phases of the cobalt succinate system to be produced from identical species reacting under different synthesis conditions was initially recognised when the dependence of the structure of the cobalt succinates on the temperature of the synthesis reaction was investigated by Forster *et al*.[78] in a study carried out with the aim of clarifying the reaction conditions that lead to the formation of distinct cobalt succinates structures. Prior to this, four cobalt succinate structures had been produced, but these structures had been synthesised from different reactants and stoichiometries, and at different pH values. In their study, Forster *et al*.[78] demonstrated that, from a single starting mixture, all four previously synthesised phases, as well as a fifth, could be produced by altering only the temperature of the synthesis reaction. Further to this, in a more comprehensive study by Forster *et al*.,[79] high-throughput experimentation was used to elucidate the precise reaction conditions by which each phase could be formed. By investigating the roles of temperature, concentration, pH and reaction time in determining the structure formed, and finding temperature and reactant ratio to be the more important reaction variables, this high-throughput work enabled the synthesis of four of the five previously reported phases and an additional two new phases. Thus, high-throughput experimental synthesis allowed the generation of a crystallisation diagram

delineating the temperatures and the reactant ratios at which the different cobalt succinate phases form, as shown in Figure 2.6.

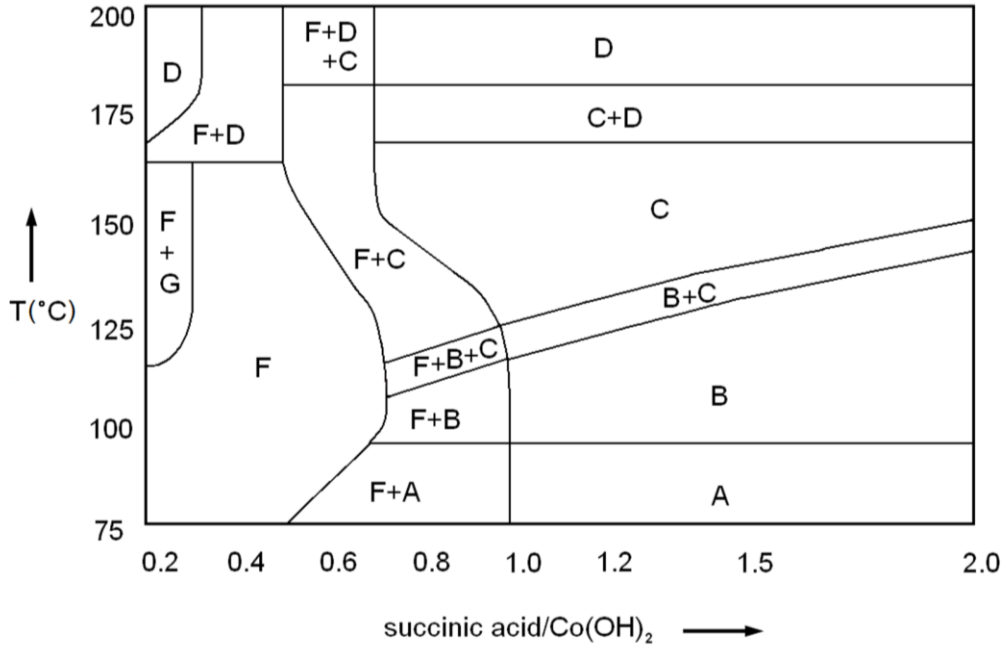

Figure 2.6: The cobalt succinate system crystallisation phase diagram in terms of reactant ratio and temperature. The letters in the phase diagram correspond to the phases shown in Table 2.1. This figure was adapted from ref. [79].

In both studies by Forster and co-workers, the cobalt succinates were formed by reacting cobalt hydroxide ($Co(OH)_2$) and the dicarboxylic acid, succinic acid ($HO_2C(CH_2)_2CO_2H$), with water as the solvent. The organic linker was provided by the succinic acid, and cobalt hydroxide provided the metal entity for the synthesis of the cobalt succinates. The initial study carried out to determine the effect of temperature involved the utilisation of the reactants in a 1:1 molar ratio. Between the two works, the phases of cobalt succinate were produced via experimental syntheses with cobalt initially at such concentration as to be present in the solvent in a molar ratio of $Co(OH)_2$:H_2O = 1:110. The formulae of the cobalt succinate phases are given in Table 2.1, and Table 2.2 gives, for Phases A-F, the experimental densities and the coordination (in terms of H_2O/Co^{2+}) in the structures. Whilst, by virtue of being different in colour to the other phases, Phase G can be easily distinguished from the other cobalt succinate structures, due to the difficulty of obtaining single crystals or phase-pure powder samples, the exact structure of Phase G has not been

established. Therefore, details relating to the experimental structure of this material are not included. Due to the intricacies of the structures of the cobalt succinate materials, for reasons of brevity, information relating to all of the bond lengths and bond angles in the structures is not included. This data can be found in the relevant references indicated in Table 2.1, however. For the same reason, because there are 22 crystallographically independent cobalt atoms in Phases A-F, the specific arrangements of ligands (and coordinated solvent molecules) around all differently environed cobalt atoms in each structure are not illustrated in this thesis. As regards the cobalt-coordination data given in Table 2.2 the first number refers to the total water content of the structure and the second (shown in parentheses) relates to the number of water molecules coordinated to the cobalt atoms in the structure.

Table 2.1: Formulae of the cobalt succinate phases.

Code	Phase	Reference
A	$[Co(H_2O)_4(C_4H_4O_4)]$	80
B	$[Co(H_2O)_2(C_4H_4O_4)]$	78
C	$[Co_4(H_2O)_2(OH)_2(C_4H_4O_4)_3]\cdot 2H_2O$	81
D	$[Co_6(OH)_2(C_4H_4O_4)_5]\cdot 2H_2O$	82
E	$[Co_5(OH)_2(C_4H_4O_4)_4]$	83
F	$[Co_7(H_2O)_3(OH)_6(C_4H_4O_4)_4]\cdot 7H_2O$	84
G	$[Co_7(H_2O)_2(OH)_{12}(C_4H_4O_4)]$	85

Table 2.2: The density and coordination around the cobalts atom in each of the cobalt succinate phases.

Code	Density (g/cm^3)	Coordination (H_2O/Co^{2+})
A	1.945	4(4)
B	1.926	2(2)
C	2.085	1(½)
D	2.197	⅓(0)
E	2.337	0(0)
F	2.165	$^{10}/_{7}(^{3}/_{7})$

The structures of Phases A-F of the cobalt succinates are shown in Figure 2.7. The structure of Phase A is one-dimensional and is composed of chains of cobalt atoms linked by succinate ions. The carboxylic group in the succinate ion is monodentate, and therefore presents a free oxygen atom, which is involved in hydrogen boding with adjacent cobalt-succinate chains. Each cobalt atom in Phase A is octahedrally coordinated to two succinate ions and four water molecules.[80] The structure of Phase B is similar to that of Phase A. However, instead of the isolated cobalt atoms linked by succinate ions that constitute Phase A, Phase B consists of clusters of three octahedrally coordinated cobalt atoms connected by succinate ions to form one-dimensional chains. In this structure, in the clusters of three cobalt atoms the two outer cobalt atoms are coordinated by three water molecules and three oxygen atoms belonging to succinate ions. The central cobalt atom in each cluster is fully coordinated by oxygen atoms from succinate ions.[78] Phase C features a two-dimensional network of edge-sharing cobalt (CoO_6) octahedra. The connection of the cobalt polyhedra forms 14-membered ring windows, which are stabilised by succinate ions, of which each is linked to four cobalt atoms. These rings form layers (as shown in Figure 2.7c) of which a stacked succession connected by hydrogen bonds produces the hydrated channels associated with this phase. This phase also differs from Phases A and B in that the cobalt atoms are coordinated by hydroxyl groups in addition to water and succinate ions.[81] The structure of Phase D constitutes a three-dimensional open network that consists of layers connected by succinate molecular pillars. The layers are constructed from edge-sharing cobalt polyhedra, which are interconnected to produce 14-membered, rhombic-shaped rings. The layers in the structure are consolidated by both hydroxyl groups and succinate ions but, in contrast with Phases A-C, no water molecules are coordinated to the cobalt atoms. In this phase (unlike Phase C in which adjacent layers are linked via weak interactions) successive layers are covalently connected by succinate groups, producing hourglass-shaped channels.[82]

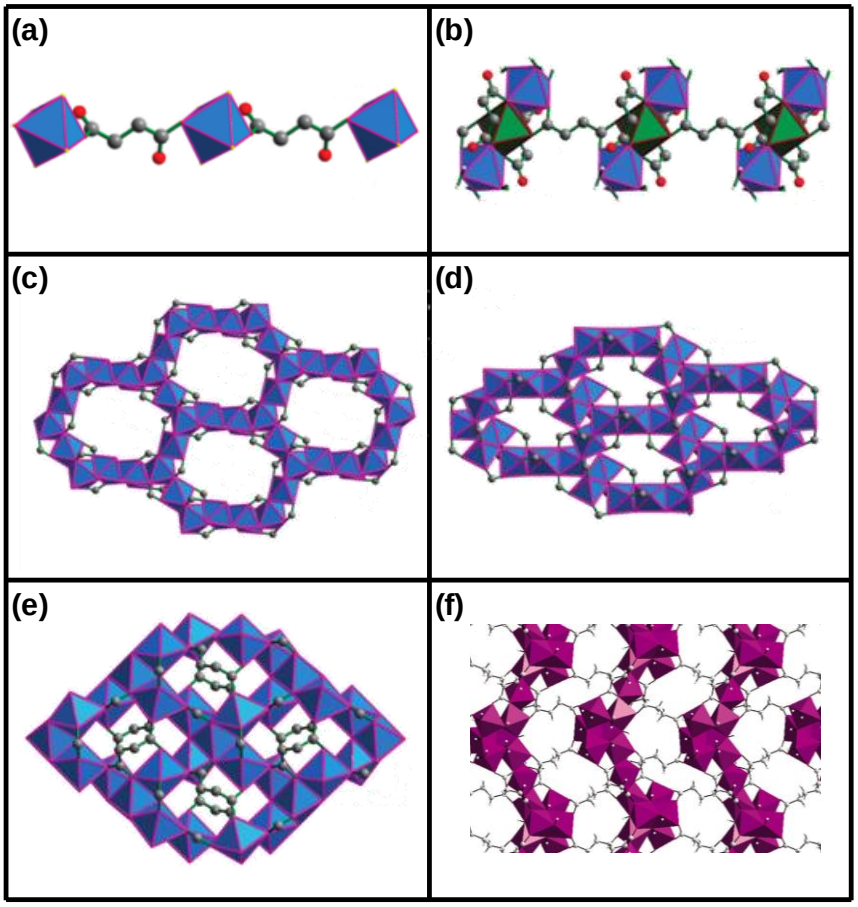

Figure 2.7: The structures of the cobalt succinate phases where (a) shows one of the linear chains that associate by hydrogen bonding to form Phase A, (b) shows the structure of Phase B, (c) shows a representation of one layer of the structure of Phase C, (d) shows a single layer of the structure of Phase D, (e) gives a representation of one layer of the structure of Phase E, and (f) shows a single layer of Phase F. In (a) – (e) the blue octahedra represent CoO_6 centres in the structures. The green octahedra in (b) represent the CoO_6 centres that are distinct from those represented by blue octahedra in Phase B. In (f) the CoO_6 centres are represented by purple octahedra. In all images, the carbon atoms in the structures are represented by grey spheres, the hydrogen atoms are represented by white spheres and the oxygen atoms are represented by red spheres. The images included in this figure were adapted with permission from ref. [86] Copyright 2008 American Chemical Society.

Phase E exists as a three-dimensional framework. This framework consists of infinite two-dimensional layers of edge-sharing cobalt octahedra that are connected by succinate ions. Like Phase D, the octahedral coordinations of the cobalt atoms in this structure arise from bonds with succinate ions and hydroxyl groups. Within the layers in this structure, the cobalt octahedra form 12-membered, lozenge-shaped cavities in which are located the alkyl chains of succinate ions. As with Phase D, in Phase E the adjacent layers are connected covalently by succinate pillars.[83] The

structure of Phase F consists of one-dimensional, covalently connected metal-oxygen-metal chains of CoO_6 octahedra that are joined by succinate-ion ligands to form an infinite sheet. Within the sheets, the cobalt coordination arises from connections with hydroxyl groups, succinate ions and water molecules. The sheets in this structure are linked by hydrogen bonding between both constituents of the sheets and unbound water molecules.[84]

Thus, the cobalt succinates are a well-characterised family of transition metal dicarboxylates, with the currently reported structures incorporating only water and/or hydroxide as additional components. In general, as regards the structures of the cobalt succinate phases, with increasing temperature there is a tendency for fewer water molecules to coordinate to each cobalt ion. Also, with increased temperature, the structures exhibit both increased edge-sharing connectivity and higher coordination numbers for carboxylate groups, and, in phases formed above 100 ^{o}C, hydroxide ions are incorporated.[78] The development of a simulation method for the synthesis of MOFs using the cobalt succinates as the test system required accurate representation of the reactants involved in the synthesis of the MOF structures. This necessitates knowledge of the chemistry of the cobalt succinate system.

2.5.1 System Chemistry

Of primary concern when developing a simulation method that attempts to predict the cobalt succinate structure synthesised under specified reaction conditions is that the entities modelled in the simulation are representative of those present in the equivalent experimental system. Thus, knowledge of the behaviour of the starting materials (in this case, succinic acid and cobalt hydroxide) with respect to one another in the specified solvent (water) was required.

In relation to the behaviour of the initial compounds in the solvent, cobalt hydroxide, which constitutes the initial compound providing the metal source in the cobalt succinate phases, is of such low solubility in water as to be considered insoluble. Therefore, for the dissociation of cobalt hydroxide in water, as given by the solubility

equilibrium shown in Equation 2.19, the equilibrium lies to the very far left (i.e., favouring cobalt hydroxide).

$$Co(OH)_2 \overset{H_2O}{\rightleftharpoons} Co(OH)^+(aq) + OH^-(aq) \rightleftharpoons Co^{2+}(aq) + 2OH^-(aq) \qquad \textbf{2.19}$$

Succinic acid, which provides the organic linker in the cobalt succinate materials, is a weak acid and, in water, exists in equilibrium with the monoprotonated and nonprotonated succinate ions. The equilibrium lies to the left (i.e., favouring succinic acid). The succinic acid dissociation is given by Equation 2.20:

$$HO_2C(CH_2)_2CO_2H \rightleftharpoons HO_2C(CH_2)_2COO^- + H^+ \rightleftharpoons {}^-OOC(CH_2)_2COO^- + 2H^+ \qquad \textbf{2.20}$$

In the preparation of cobalt succinates, when cobalt hydroxide and succinic acid are added to the solvent, the dissociated Co^{2+} ions react with the succinate ions in solution, as shown in the complexation reaction given in Equation 2.21.

$${}^-OOC(CH_2)_2COO^- + Co^{2+} \rightleftharpoons CoO_2C(CH_2)_2CO_2 \qquad \textbf{2.21}$$

The Co^{2+} ion is very labile, with the effect that complexation reactions involving Co^{2+} usually occur very quickly. As the Co^{2+} ions and succinate ions in solution undergo complexation, the equilibria represented in Equations 2.19 and 2.20 will move to the right, with the effect that more succinic acid will be ionised and more cobalt hydroxide will dissociate. Since syntheses of MOFs typically take place in solution (as opposed to at a surface), the formation of cobalt succinates was assumed to occur in solution rather than at the surface of the undissolved cobalt hydroxide. Thus, in syntheses of cobalt succinate materials, the components involved in self-assembly were assumed to be Co^{2+} ions and succinate ions. In this work, for simplicity, only the species involved in the self-assembly process, which, experimentally, yields the desired product, were considered throughout the

simulations. Consequently, changes in the starting components arising from their combined immersion in the solvent were not represented in the simulations. Therefore, in simulating the self-assembly process by which the cobalt succinate materials form, only Co^{2+} ions and succinate ions were represented as solutes. For further simplification, the self-assembly process was assumed to involve only the nonprotonated form of the succinate ion. The structure of the nonprotonated succinate ion is given in Figure 2.8.

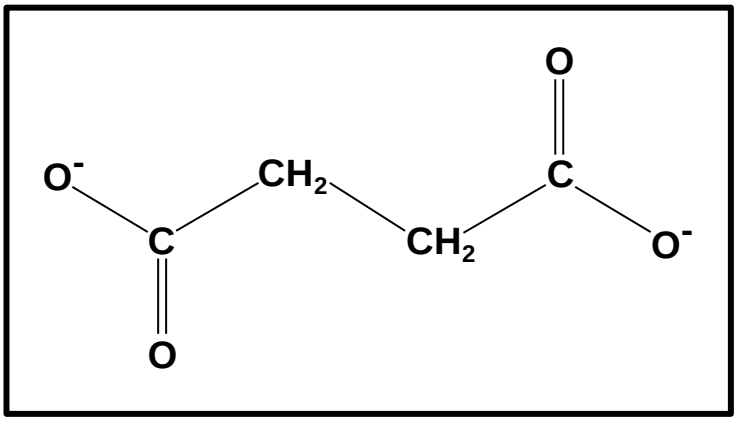

Figure 2.8: Nonprotonated succinate ion

In addition to the potential of the cobalt succinate materials to provide a thorough test of the capability of a method established for simulation of the synthesis of MOFs, the formation of cobalt succinates by reacting Co^{2+} ions and succinate ions in water presented several advantages with respect to the modelling of such systems. With only two reactant species involved in the formation of cobalt succinates, the number of molecules to be modelled was kept to a minimum. Also, in relation to the minimisation of substances to be modelled, the physical reaction does not yield any salt by-products. This has the added benefit of preventing any difficulties arising from the necessity to consider the role of different bases, additional phases that may incorporate cations from the bases, and the influence on the reaction of the salt by-product concentration.

2.5.2 Cobalt Succinate System – Molecular Potentials

For the initial system studied, three types of molecule were required to simulate the experimental situation: (i) the linker species, the succinate ion, (ii) the vertex species, the Co^{2+} ion, and (iii) the solvent molecule, water. The ability of a simulation of the self-assembly of the reactant species to realistically replicate the experimental behaviour of the reactants is dependent on the accuracy of the molecular potentials used to represent the interactions between species in the system. As regards modelling the succinate ion, the OPLS (optimised potentials for liquid simulations) molecular potentials, using a partially united-atom model (OPLS-UA), have been shown to correctly reproduce thermodynamic and structural data (in terms of the interaction energies and geometries predicted by *ab initio* calculations, and experimental hydration enthalpies and primary hydration numbers) for dilute solutions of the monocarboxylate ions formate ($HCOO^-$) and acetate (CH_3COO^-).[87] The applicability of these potentials for the representation of difunctional carboxylate ions such as the succinate ion has not been assessed. However, using the equivalent all-atom (OPLS-AA) potential functions, and employing parameters derived for monofunctional molecules, the simulated behaviour of both the difunctional succinic acid and the corresponding monoanion in water have shown excellent agreement with experimental results.[88]

Therefore, by analogy, the OPLS-UA parameters derived for monofunctional carboxylate ions were applied to the succinate ion. In this representation each atomic nucleus is represented by an individual interaction site except CH_n groups, which are treated as united atoms centred on the carbon atom. Where appropriate, the OPLS-UA parameters for carboxylate ions[87] were adopted from previously determined OPLS-UA values for hydrocarbons[89] and amides.[90, 91] In addition, the OPLS-UA parameters for carboxylate ions specify that the C-O bond length and O-C-O bond angle have values of 1.231 Å and 131°, respectively.[87] However, the OPLS-UA carboxylate ion parameter set does not give values suitable for the CH_2 group in the succinate ion. Therefore, to ensure correspondence, OPLS-UA parameters developed for hydrocarbons were utilised and the CH_2 united atom was represented

using LJ parameters developed for the sp^3 hybridised CH_2 group in n-butane. Similarly, the CH_2-CH_2 bond length was taken as that between two sp^3 hybridised carbons in a hydrocarbon molecule. Furthermore, following the approach used in the aforementioned study of the sampling of the monoprotonated succinate ion,[88] the CH_2-C-O bond angle in the succinate ion representation was specified according to the corresponding OPLS-UA parameters defined for amides. The dihedral angles in the succinate ion were specified according to the OPLS-UA parameters developed for hydrocarbons[89] and, where the OPLS-UA parameters were deficient, the appropriate OPLS-AA values.[92]

In the model of the succinate ion the interaction sites were assigned charges, corresponding to those specified in the OPLS-UA carboxylate ion parameter set,[87] such that the overall charge on the ion was -2, thus maintaining a physically realistic depiction of the system. The geometry of the interaction sites in the succinate ion is described in terms of the bond lengths, bond angles and dihedral angles given, respectively, in Tables 2.3, 2.4 and 2.5. The LJ parameters and the charges associated with the interaction sites in the succinate ion are given in Table 2.6. As regards intramolecular interactions, the OPLS-UA force field involves fixed bond lengths and angles but, for hydrocarbons, alcohols and amides, parameters describing torsional motion have been developed. However, the aforementioned study regarding the conformational sampling of difunctional succinic acid required re-parameterisation of the OPLS-AA torsional parameters developed for monofunctional carboxylic acids,[88] indicating that the parameters describing torsional motion are not easily transferable. Thus, since the OPLS-UA torsional parameters were not developed in relation to difunctional carboxylate ions, rather than include incorrect torsional motion, the succinate ion was modelled as a rigid species, with intramolecular interactions considered within the succinate ion only between pairs of non-bonded atoms that are separated by more than four bonds.

Table 2.3: The bond lengths between interaction sites in the succinate ion model. The reference relates to the initial publication of the corresponding OPLS-UA parameter.

Geometry	Bond Length (Å)	Reference
CH_2-CH_2	1.53	[89]
C-CH_2	1.522	[90]
C-O	1.231	[87]

Table 2.4: The bond angles in the succinate ion model. The reference relates to the initial publication of the corresponding OPLS-UA parameter.

Geometry	Bond Angle (°)	Reference
C-CH_2-CH_2	112.0	[89]
CH_2-C-O	120.5	[90]
O-C-O	131.0	[87]

Table 2.5: The dihedral angles in the succinate ion model. The reference relates to the initial publication of the corresponding parameter.

Geometry	Dihedral Angle (°)	Reference
C-CH_2-CH_2-C	180.0	[89]
CH_2-CH_2-C-O	180.1	[92]

Table 2.6: LJ parameters and charges associated with interaction sites in the model of the succinate ion. The reference relates to the initial publication of the corresponding OPLS-UA parameter.

Site	σ (Å)	ε/k_B (K)	q(eu)	Reference
O	2.960	105.676	-0.8	[87]
C	3.750	52.838	0.7	[87]
CH_2	3.905	59.380	-0.1	[87, 90]

The Co^{2+} ion was modelled, as a single interaction site, using a force field developed for divalent metal cations in water.[93] Through adjustment of the LJ parameters describing the ion-water interaction in such a way that the hydration free energy of the ion can be replicated, this force field can capture the microscopic solvent molecular effects around the cation. The parameters for the Co^{2+} ion were developed for use in conjunction with proteins. However, the constituents of the organic linker

are sufficiently similar to those of proteins that the force field can reasonably be used to represent the intermolecular interactions calculated by pairwise summation. As with the succinate ion, the cobalt ion was assigned the natural charge value of +2. The LJ parameters and the charge associated with the Co^{2+} ion are shown in Table 2.7.

Table 2.7: Parameters associated with the Co^{2+} ion model

Site	σ (Å)	ε/k_B (K)	q(eu)	Reference
Co	2.186	14.392	2.0	[93]

Interactions with the solvent play a critical role in the simulation of a condensed system. Thus, the molecules constituting the solvent within which the Co^{2+} ions and the succinate ions are immersed must be accurately represented. The parameters for the Co^{2+} were obtained in conjunction with the use of the TIP3P[94] water model, which has three interaction sites, corresponding to the three atoms of the water molecule. All OPLS models, such as the OPLS-UA force field employed in the description of the succinate ion, have been developed to be compatible with the TIP4P[94] representation of water molecules. The TIP4P water model is a four-site representation, in which an additional charge is placed at a fictive site along the bisector of the hydrogen sites. Both the TIP3P and TIP4P models involve rigid representations of the water molecule, and so neither model includes intramolecular interactions. The arrangements of the interaction sites in the TIP3P and TIP4P water models are shown in Figure 2.9. Despite their inherent differences, the TIP3P model yields very similar results to the TIP4P model to the extent that the two models can, for most purposes, be considered interchangeable.[95] Thus, in this work, the TIP3P water was utilised in the representation of explicit water molecules because the TIP4P model does not offer a sufficient improvement in accuracy to justify the additional computational cost resulting from the higher number of interaction sites. The structural parameters associated with the TIP3P water model are given in Table 2.8, and the LJ parameters and charges relating to the interaction sites in the TIP3P model are shown in Table 2.9.

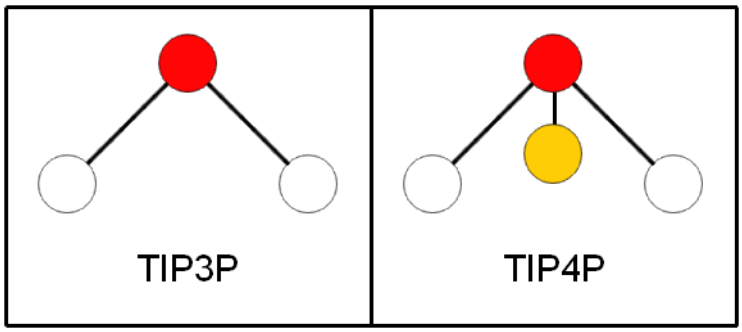

Figure 2.9: Structures of the TIP3P and TIP4P representations of water molecules, in which the red circles denote oxygen atoms, the white circles denote hydrogen atoms, and the yellow circle represents a dummy site.

Table 2.8: The bond length and angle involved in the TIP3P water model.[94]

Geometry	Bond Length (Å) / Angle (°)
O-H	0.9572
H-O-H	104.52

Table 2.9: LJ parameters and charges associated with interaction sites in the TIP3P water model.[94]

Site	σ (Å)	ε/k_B (K)	q(eu)
O	3.151	76.526	-0.834
H	0.0	0.0	0.417

In order to examine the suitability of the outlined molecular potentials for representing the reactants and the solvent involved in the synthesis of cobalt succinates, simulations of the self-assembly of Co^{2+} ions and succinate ions in water were carried out, as described in Section 2.6.

2.6 Explicit-Solvent Self-Assembly Simulation

The self-assembly of Co^{2+} ions and succinate ions in explicitly-modelled water was represented via MC simulations in the canonical ensemble. To examine the suitability of the simulation methods and the molecular potentials employed, the explicit-solvent self-assembly simulation was executed at a temperature and reactant ratio corresponding to those for which Phase A of the cobalt succinates forms

experimentally. Thus, the simulation was carried out at a temperature of 75 ^{o}C and with a succinate:Co^{2+} ratio of 1.2:1. Initially, the least computationally intensive size of system was studied, and a simulation of the smallest possible system containing entire reactant ions was carried out. Thus, corresponding to the succinate:Co^{2+} ratio of 1.2:1, the smallest simulation cell included six succinate ions and five Co^{2+} ions. The ions were modelled using the parameters detailed in Section 2.5.2. In compliance with the experimental Co^{2+}:H_2O ratio of 1:110, the solvent in the simulation cell was represented by 550 TIP3P water molecules. The reactant species and the solvent were contained within a cubic simulation box, the size of which was dictated by the requirement for the density to be equivalent to that of the experimental system at the specified temperature. For this system, the dimensions of the simulation box were 26.43 Å.

The starting configuration for the explicit-solvent self-assembly simulation was generated by means of grand canonical Monte Carlo (GCMC) simulations.[56] In addition to the translation and rotation moves by which different configurations are sampled in the canonical ensemble, GCMC simulations also involve the insertion and deletion of molecules. By terminating GCMC simulations when the simulation box contains a specified number of molecules, a configuration comprising the requisite numbers of each species, with no overlap between molecules, can be produced. The starting configuration for the self-assembly simulation was created by firstly inserting the required number of water molecules in the simulation cell. Following equilibration, by means of a canonical MC simulation, of the system containing the water molecules, the correct numbers of cobalt and succinate ions were added to the simulation cell. The GCMC simulations and the equilibration simulation, as with the self-assembly simulations in the canonical ensemble, were executed using the simulation code Music.[96] In all of the simulations performed the electrostatic interactions between species were calculated using Ewald summation and van der Waals interactions were represented using the LJ potential. The LJ potentials were cut-off at 12.0 Å, which is greater than three times the largest LJ sigma value associated with the species involved in the simulation and, thus, represents a distance at which the interaction energies are negligible.

From the starting configuration, the self-assembly of Co^{2+} and succinate ions in water was simulated for 4×10^9 steps, taking approximately six weeks. The starting configuration, and the final configuration, viewed along the x-, y- and z- axes are shown in Figure 2.10.

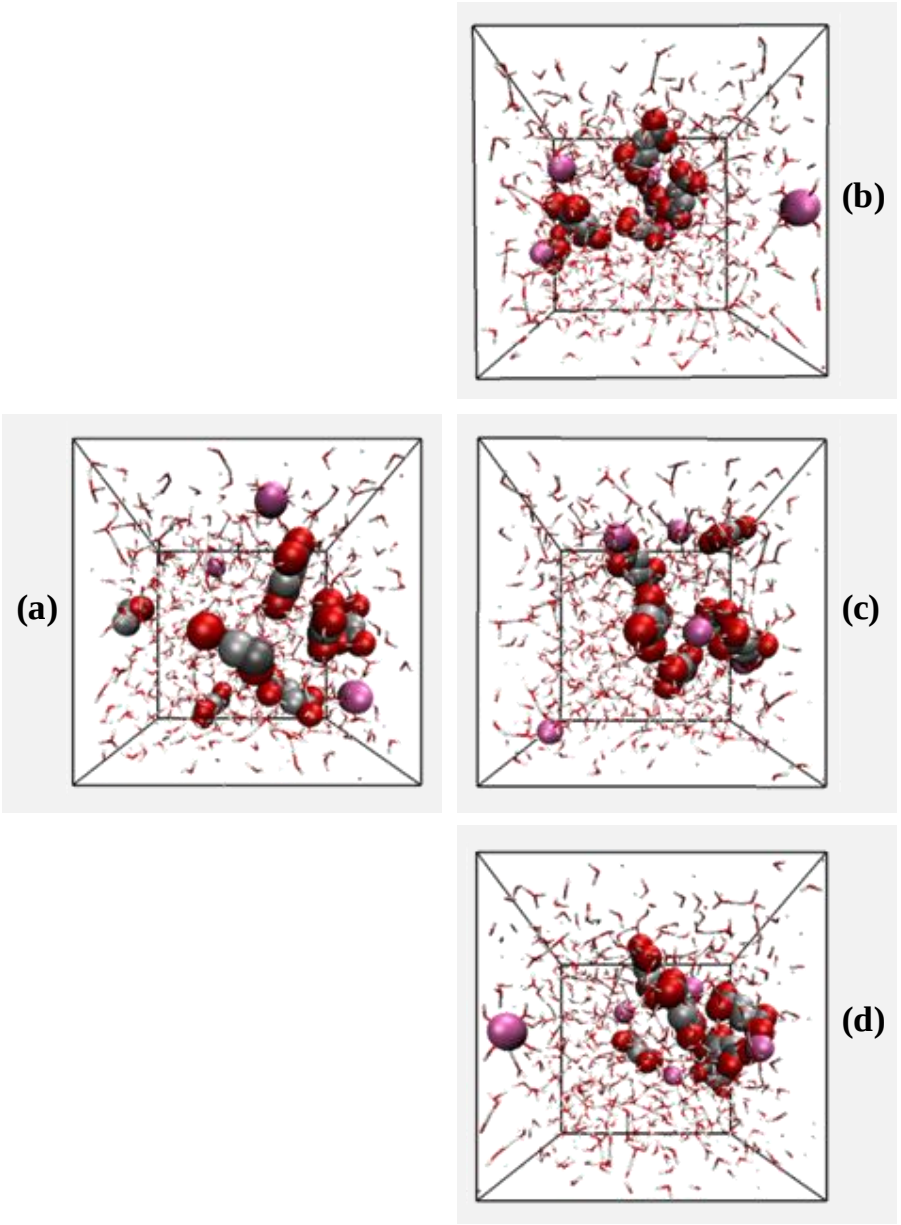

Figure 2.10: Configurations sampled in the explicit-solvent self-assembly simulation where (a) represents the starting configuration, and (b), (c), and (d) represent, respectively, the final configuration viewed along the x-, y- and z-axes. In these images the cobalt ion is represented by a pink sphere and in the succinate ion oxygen is represented by a red sphere, carbon is represented by a dark grey sphere, and methylene is represented by a light grey sphere. The water molecules in the simulation cell are represented in 'stick' form, with the oxygen and hydrogen atoms denoted, respectively, by red and white regions.

The images of the final configuration shown in Figure 2.10 demonstrate that, after a large number of steps, the Co^{2+} and succinate ions have undergone a degree of self-assembly, and, within the simulation cell, the ions are located over a smaller region than in the starting configuration. However, there is no clear arrangement of the ions, and a number of ions remain isolated and surrounded by layers of water. Comparison of the average distances between the centres of mass of the ions in, respectively, all pairs of succinate-succinate, Co^{2+}-succinate, and Co^{2+}-Co^{2+} indicates, as shown in Figure 2.11, that there was minimal variation in the relative positions of the ions throughout the simulation. Indeed, throughout the extent of the explicit-solvent simulation, the average separation distances associated with the pairs of ions varied by less than 10 % of the initial value.

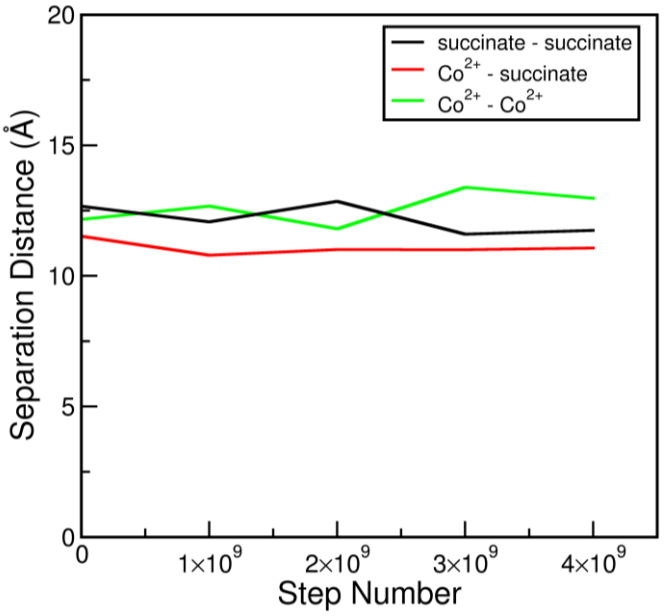

Figure 2.11: Variation, with step number, in the average separation distance between the centres of mass of ions involved in pairs of succinate-succinate, Co^{2+}-succinate, and Co^{2+}-Co^{2+}.

The lack of variation in the relative positions of the ions in the explicit-solvent simulation indicates that there were only marginal differences between successive configurations sampled throughout the simulation. This is a consequence of the presence of the water molecules in the simulation box, such that the movement of the ions was restricted to very small alterations in position. Large shifts in the positions

of the ions would result in overlap with the water molecules and were, thus, unfavourable. Consequently, even for the small system studied, significantly more simulation steps were required to achieve an appreciable extent of self-assembly.

To be practical, a method developed for simulating the synthesis of MOFs must be capable of predicting the behaviour of larger systems than the one studied and within reasonable simulation time. Therefore, the computational cost associated with the use of an explicit-solvent representation is not feasible. In order to reduce the computational demands of the self-assembly simulation, the use of an implicit-solvent approach is required.

CHAPTER 3

Calculation of the potential of mean force between ionic species

Molecular simulations in which solvent molecules are modelled explicitly as discrete entities provide the most accurate and detailed description of the system under investigation. However, simulations with explicitly modelled solvent molecules have high numbers of degrees of freedom, arising from all of the modes of motion of all simulated molecules, and are therefore very computationally intensive, requiring prohibitively long computational times to obtain meaningful results with regard to the properties of the solutes. Thus, there is great advantage in accounting for the influence of the solvent environment in a manner that is both computationally efficient and physically realistic. Implicit-solvent approaches, in which the effects of the solvent are approximately described without including an explicit representation of the solvent molecules, reduce the time for each interaction calculation and allow much faster sampling of solute configurations and phase space.

Several implicit-solvent techniques have been developed to allow the exclusion of large numbers of solvent molecules from simulations. Whilst an outline description of these methods follows, detailed reviews can be found elsewhere.[97-100] The primary concern when implementing an implicit-solvent method is that the simulation accurately represents the necessary solvent effects. Therefore, as is appropriate to the extent of the range of solvated systems of interest, implicit-solvent representations vary in the level of detail they feature, and, by association, the computational cost. At the opposite end of the spectrum of solvent detail from explicit methods, the most computationally efficient implicit-solvent approaches are based on continuum electrostatic approximations, in which the solvent environment is represented as an infinite, featureless dielectric medium. However, the applicability of continuum methods is limited by the loss of atomistic detail and the associated lack of accuracy.[97] Continuum electrostatic representations are often combined with solvent-accessible surface area (SASA) models,[101, 102] which approximate the nonpolar contribution to the solute-solvent interaction by assuming that the solvent-induced interaction is proportional to the exposed surface area of the solute molecule. However, SASA-based approaches are liable to inaccuracy as a result of the failure to account for the attractive van der Waals interactions between solvent and solute atoms.[103] Other implicit-solvent methods include hydration shell models,[104, 105] in which the solvation energy of a solute is assumed to be proportional to the solvent-accessible volume of a defined solvation shell, and statistical mechanical theories based on integral equations.[106, 107] The ability of hydration shell methods to represent solvent effects are limited by the exclusion of molecular-scale detail.[108] Techniques based on integral equations do consider the solvent molecular structure but there is some uncertainty as to their precision and accuracy.[109] Thus, there are advantages and disadvantages associated with all of the different implicit-solvent methods. In simulations of self-assembly, to generate realistic configuration of the solute species, the implicit-solvent method must combine computational efficiency with the ability to include sufficient solvent structural detail. Therefore, in this work, the solvent was modelled implicitly by incorporating the effects of the solvent using a ‘potential-of-mean-force’ (PMF) approach to represent the

interactions between the self-assembling species. The PMF methodology is described in Section 3.1.

3.1 Potential-of-Mean-Force Theory

First introduced by Kirkwood[110] in 1935, the PMF along a coordinate of the system, defined as the reaction coordinate, is a measure of the relative free energy of the system as a function of the selected coordinate. With the reaction coordinate defined as the distance between two species that are immersed in solvent, the PMF represents the free-energy profile along the separation distance. By averaging over all the degrees of freedom of the system other than the reaction coordinate (i.e., the relative orientations of the solute species and the conformations of the surrounding solvent), a PMF in terms of separation distance, $W(r)$, can be obtained. Such a PMF provides a fundamental measure of the reversible work required to bring together the solutes from infinitely far apart to a separation r, in the solvent. Therefore, the PMF encompasses some of the solvent-mediated behaviour in an effective solute-solute interaction potential, which replaces the explicit treatment of the solvent-solvent and solvent-solute interactions. Thus, the use of PMFs allows simulations to be carried out such that only the solute molecules are explicitly represented. The effect of the utilisation of PMFs, as a means of representing the effects of the solvent molecules in a simulation, is illustrated in Figure 3.1.

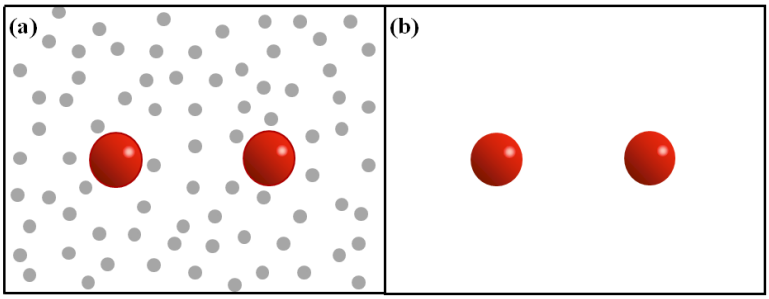

Figure 3.1: For a pair of solute molecules, the interactions considered between the species present with an explicit-solvent representation, as shown in (a), can be replaced by a PMF between the solutes so that only a single interaction potential between the solutes is required with the implicit-solvent representation shown in (b). In these images grey spheres represent the solvent molecules and red spheres represent the solute species.

The free energy at a particular separation, $W(r)$, is directly related to the equilibrium probability, $P(r)$, of finding the solute species a certain distance apart in the solvent:

$$W(r) = -\mathrm{k_B}\, T \ln P(r) \qquad \textbf{3.1}$$

where T is the absolute temperature of the system and $\mathrm{k_B}$ is the Boltzmann constant. In order to calculate the correct probability distribution along the reaction coordinate, all regions of the reaction coordinate must be adequately sampled. In principle, standard Monte Carlo (MC) or Molecular Dynamics (MD) techniques can be employed to obtain the distribution of probabilities with which each configuration along a reaction coordinate is sampled. However, the logarithmic nature of the relationship between $W(r)$ and $P(r)$ means that for regions of the reaction coordinate where the PMF exceeds a magnitude of a few times $\mathrm{k_B}T$, the probability of occurrence is very low. Therefore, in a typical MC or MD simulation, to sample all possible conformations of interest along the reaction coordinate and obtain statistically meaningful results would require unfeasibly long simulation times. Several approaches have been developed to circumvent the sampling difficulties and calculate PMFs along reaction coordinates via molecular simulations. Among them, free energy perturbation (FEP),[111-114] thermodynamic integration (TI),[112-115] and umbrella sampling[116] combined with the weighted histogram analysis method (WHAM)[117, 118] are frequently used for calculating PMFs. Determination of a free-energy profile using the FEP method[119, 120] involves calculation of the sum of the differences in the free energy between a succession of incremental perturbations representing intermediate states between the initial and final points of interest along the reaction coordinate, whereas the TI technique[121] allows computation of a PMF by integrating the derivative of the free energy along the reaction coordinate path. Whilst both FEP and TI approaches have been successfully applied to determine PMFs, the FEP method has been shown to cause instability and hysteresis and, consequently, poor convergence, and TI has been found to propagate errors.[122] Thus, in this work, to determine the PMF between a pair of solutes, as an efficient method for sampling low-probability events, umbrella sampling was employed to guarantee

exploration of all regions of the relevant reaction coordinate. The umbrella sampling technique is outlined in Section 3.1.1.

3.1.1 Umbrella Sampling

The umbrella sampling method, first detailed by Torrie and Valleau,[123, 124] involves simulating the microscopic system under investigation in the presence of an artificial biasing potential, which is added to the potential energy. As shown in Equation 3.2, the overall energy is augmented by addition of the biasing potential:

$$U'(r) = U(r) + V(r) \qquad \textbf{3.2}$$

where $U(r)$ represents the total energy of the unbiased simulation, $V(r)$ is the biasing potential, and $U'(r)$ is the total energy of the biased simulation. The purpose of the biasing potential is to force the simulation to explore regions of conformational space that are inadequately sampled in unbiased simulations. Thus, the efficiency of umbrella sampling is greatly influenced by the selection of the biasing potential. Ideally, the biased simulation would produce a uniform distribution of conformations along the reaction coordinate. However, since this could only be achieved through the use of a biasing potential equivalent to the negative of the PMF,[125] which, clearly, has not been established prior to the start of the calculation to determine the PMF, the biasing potential must take a different form.

Therefore, in umbrella sampling, to ensure that the entire pathway of interest is sufficiently well sampled, the reaction coordinate is divided, as shown in Figure 3.2, into a number of regions (termed windows), which are simulated independently. The region of the reaction coordinate being explored in each window is simulated in the presence of a particular biasing potential which is selected to restrain the system to sample configurations within a small region of a prescribed value, which is typically the centre of the window. Thus, multiple simulations are performed with different biasing umbrella potentials that centre the sampling in overlapping regions along the reaction coordinate.

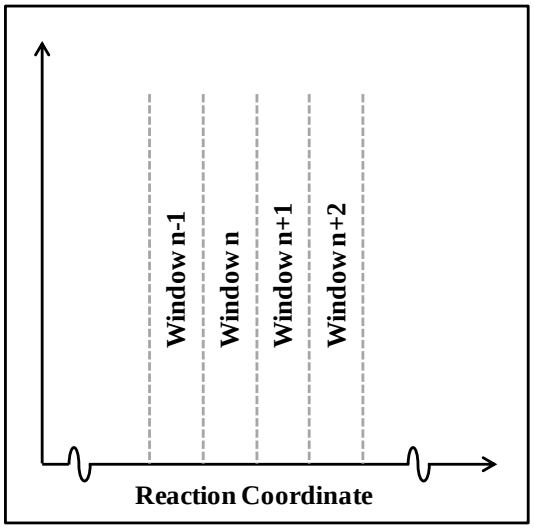

Figure 3.2: To determine a PMF, the utilisation of umbrella sampling requires that the reaction coordinate is divided into a number of windows.

The biasing potential utilised in umbrella sampling usually takes the form of a simple harmonic function, which keeps the simulation localised to a specified window:

$$V_i(r) = \frac{1}{2} K_i (r - r_i)^2 \qquad \textbf{3.3}$$

where r is the value of the reaction coordinate, and K_i and r_i are, respectively, the harmonic force constant and the central location of window i, at which the harmonic function, $V_i(r)$, is centred and around which the sampling will be enhanced. Consequently, umbrella sampling yields a biased probability distribution of the states sampled in each window along the reaction coordinate. In practice, in simulations involving the implementation of umbrella sampling, such distributions of the probabilities with which states are visited are obtained via normalisation of histograms measuring the frequency with which the states are sampled. Such histograms are essentially arrays of equally sized bins that correspond to increments of the reaction coordinate, with the bin size dictated by the requirement for the histogram to be discretised so as to achieve satisfactory resolution. Each bin in a histogram contains a number of counts, which represents the number of visits to the particular region of the reaction coordinate contained within the bin. Throughout the

simulation, counts are added to the appropriate bins in the histogram either after every step, or following a set block of steps of stipulated length.[126] From the probability distributions produced in umbrella sampling simulations, the PMF of interest can be obtained using the WHAM technique, which is described in Section 3.1.2.

3.1.2 Weighted Histogram Analysis Method

The PMF for the unbiased system in a given window, $W_i^u(r)$, is given by:

$$W_i^u(r) = -\mathrm{k_B}T \ln P_i^b(r) - V_i(r) + C_i \qquad \textbf{3.4}$$

where $P_i^b(r)$ is the biased probability distribution obtained in window *i*, $V_i(r)$ is the biasing potential applied in window *i*, and C_i is an unknown constant that represents the free energy associated with introducing the bias potential to window *i*. In order to calculate the complete PMF from the probability distributions obtained along a reaction coordinate from a series of windows in which artificial biasing potentials have been applied, the effect of the biasing potentials must be eliminated and the data from individual windows must be combined. This can be achieved using the constant-temperature WHAM, which is derived in full in the original publication by Kumar *et al.*[117] The WHAM calculates the unbiased probability distribution as a weighted sum over the data extracted from the N_W individual windows in which umbrella sampling simulations were carried out. Thus, by utilising the overlap between the biased probability distributions determined from umbrella sampling simulations in adjacent windows and without discarding any of the data from the biased simulations, the WHAM analysis involves calculation of the minimum variance estimate of the constants C_i and, by association, the unbiased probability distribution, $P(r)$, by iteratively solving Equations 3.5 and 3.6 to self-consistency.

$$P(r) = \frac{\sum_{i=1}^{N_W} N_i P_i^b(r)}{\sum_{j=1}^{N_W} N_j \exp([C_j - V_j(r)]/\mathrm{k_B}T)} \quad \textbf{3.5}$$

$$\exp[-C_i/\mathrm{k_B}T] = \int \mathrm{d}r \exp[-V_i(r)/\mathrm{k_B}T]P(r) \quad \textbf{3.6}$$

N_i is the total number of configurations sampled in window i used to construct the corresponding biased probability distribution $P_i^b(r)$. After determination of the optimal estimate of the unbiased probability distribution, the unbiased PMF can be calculated using Equation 3.1.

In the determination of unbiased PMFs by means of umbrella sampling followed by WHAM analysis there are a number of considerations relating to the associated simulation procedures. These considerations are outlined in Section 3.2.

3.2 Simulation Details

3.2.1 PMF Entity

For the cobalt succinate system, in which structures are synthesised following the reaction between the initial solutes (cobalt ions and succinate ions) in the solvent (water), implicit representation of the solvent can be achieved through the use of PMFs to represent the effective interaction between these initial solute species. However, because the PMF is determined by averaging over all degrees of freedom of the solvent and the relative orientations of the solutes, to obtain PMFs that adequately describe effective implicit-solvent potentials between the solute ions in their entireties would, as a consequence of the aspherical shape of the succinate ion, require consideration of a complicated array of solute-solute separation distances and approach angles. This would necessitate the application of the umbrella sampling pseudopotential in a number of Ångström and degree increments over the number of

windows required to calculate the PMFs. The computational requirements of this would be prohibitive. Therefore, in order to reduce both the computational and human effort involved in the data collection necessary to determine implicit-solvent effective interactions between the solute ions, PMFs were calculated individually for each pair of atoms from which the solute ions were comprised, over the separation distance of interest. Since all atoms in the solute ions can be represented as spherical entities, there is no requirement for consideration of the effects of the angle of approach when calculating PMFs between atoms.

The use of pair PMFs to represent the interactions between constituents of molecules in self-assembly simulations has been shown to be capable of successfully capturing the behaviour of the corresponding explicit-solvent atomistic simulations. In preliminary simulation studies of the self-assembly of the hydrophobic dipeptide molecule, diphenylalanine, Villa *et al.*[127] found reasonable agreement between all-atom, explicit-solvent self-assembly simulations and implicit-solvent, coarse-grained simulations with the diphenylalanine molecule represented by seven small, spherically symmetric molecules corresponding to fragments of the dipeptide and with nonbonded interactions between the fragments described by pair PMFs. Similarly, for implicit-solvent simulations of the self-assembly of flexible surfactant molecules, Shinto *et al.*[128] used PMFs between small molecules approximating, separately, the hydrophilic and hydrophobic regions of the surfactant molecule to describe the effective interactions between molecules and obtained reasonable agreement with the equivalent atomistic simulations and with experimental measurements of the ionic surfactant molecules in aqueous solution. However, in this work, the effective interactions directing self-assembly take the form of PMFs between pairs of atoms in the molecules involved in self-assembly rather than between pairs of small molecules representative of regions of the self-assembling entities. The approach employed in this work is likely to allow the incorporation of a more detailed representation of the solvent structuring in relation to the individual atoms in the solute species.

In summary, in this work, so as to ensure computational practicability, the implicit-solvent approach is based on the calculation of separation-distance-based PMFs between the constituent atoms in the solute ions that are involved in self-assembly in the cobalt succinate system. Utilisation of this approach is intended to allow the execution of implicit-solvent simulations of the self-assembly process with only the solute species explicitly represented, and the solvent-mediated effects included implicitly by the pairwise summation of PMFs between the atom constituents of the solute species. The strategy utilised in this work, such that the associated PMFs describe the solvent-influenced interaction between pairs of atoms that are represented as spherical species, eliminates any requirement for consideration of the angles of orientation of the solute species. This necessitates evaluation of PMFs in one dimension only, and minimises the computational expense of the development of the PMF-based implicit-solvent method. In addition to simplifying the PMF-calculation requirements, the benefit of such an approach is the easy transferability to application in simulations representing the synthesis of cobalt succinate materials. Such simulations, by necessity, must involve representations of the species formed in the reactions between the initial solute ions. Therefore, an implicit-solvent method involving PMFs between pairs of atoms (which are common to both initial solute species and the species produced in reactions) can be implemented much more straightforwardly than could a method that involved PMFs between pairs of the solute species in their entireties. In order to ensure that the necessary atom-pair PMFs were accurately calculated for implementation in the self-assembly simulations, as previously described, in this work PMFs were calculated by utilising umbrella sampling followed by application of the WHAM formulation. The steps involved in the calculation of a PMF between a pair of atoms are depicted in Figure 3.3.

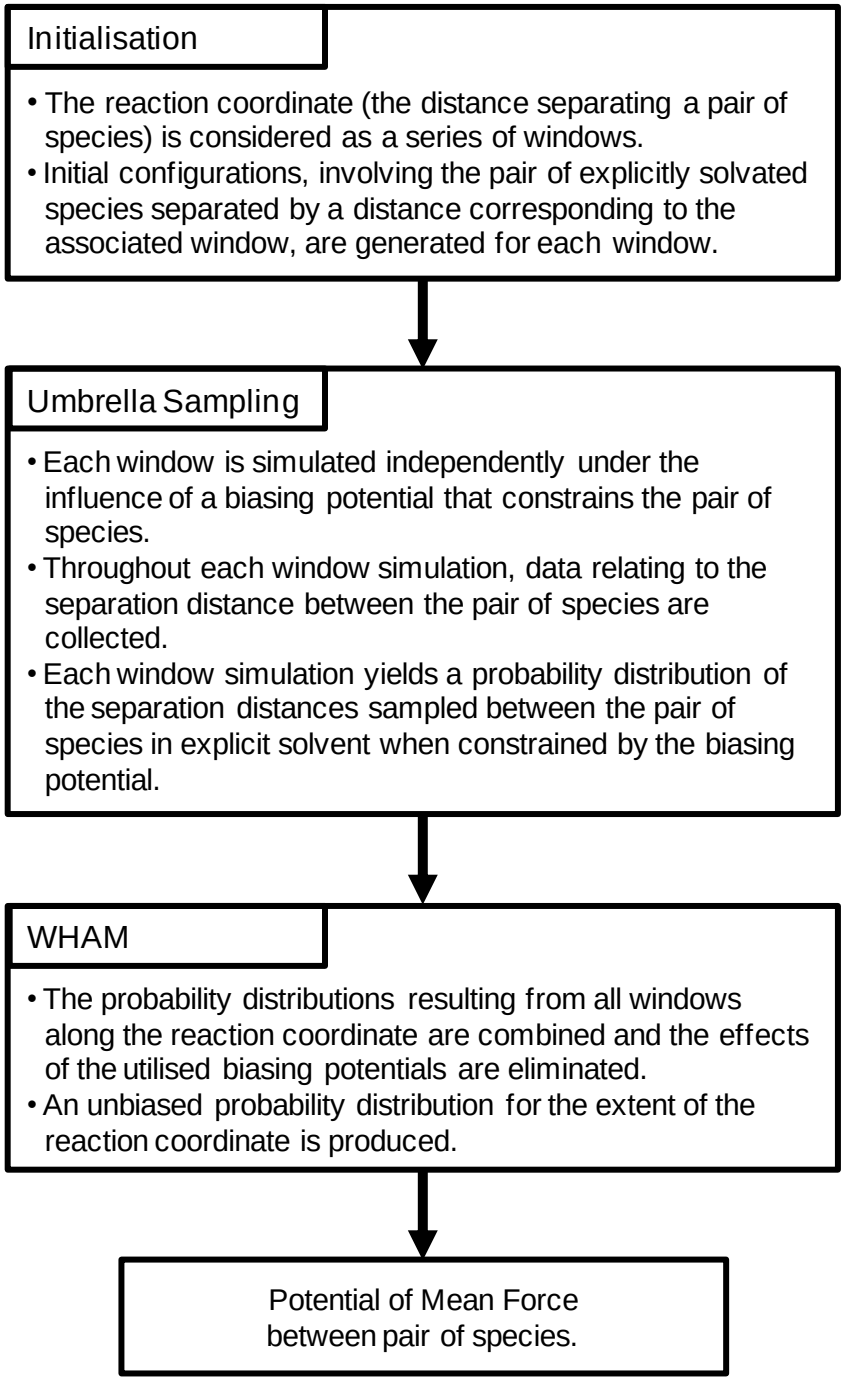

Figure 3.3: Flow chart showing the procedures involved in the calculation of a PMF that is accurately represented over all regions of the reaction coordinate. In this work, the species involved in the calculation of PMFs identify as the individual constituent atoms of the solute molecules that self-assemble.

3.2.2 Electrostatic effects

In order that the PMFs between the atom constituents of the self-assembling species can replicate the interactions between the solute ions, the atoms were modelled using parameters identical to those used when the atom existed in the solute-ion environment. Since the atoms in the overall charged solute ions are assigned individual charges, the PMFs were calculated between pairs of charged species. This required the evaluation of Coulombic interactions in the calculation of PMFs between pairs of atoms. However, when determining the PMF between a pair of atoms with associated charges, selection of a suitable method for treatment of electrostatic interactions is fundamental to the accurate computation of the free energy profile. The customary procedures for calculation of the electrostatic interactions involve either a truncation method, or the Ewald summation technique.[56] Both approaches introduce approximations into the treatment of the system and have inherent shortcomings associated with the calculation of a PMF between a pair of charged species. Whilst truncation methods, in which interactions are considered only within a sphere of specified cut-off radius, are computationally more efficient than Ewald summation, such schemes are liable to introduce errors as a consequence of the spherical cut-off even with the incorporation of one of the variety of available shifting or switching functions used to modify the Coulomb potential.[129] However, while the summation to infinity of the periodic replicas of the simulated system that characterises Ewald summation means that errors associated with truncation are not an issue, the order imposed by the periodic boundary conditions can influence the accuracy of the calculation of PMFs between charged species because the ions between which the PMF is to be determined will be repeated periodically. As a result, the charged atoms between which the PMF is calculated will interact with their periodic images,[130] and the PMF obtained will, therefore, include the effect of the periodic copies of the ions in the unit cell. In addition, for calculation of a PMF between species with unlike charges, the evaluation of electrostatic interactions using Ewald summation is further complicated because the Ewald sum converges to the electrostatic sum only for systems with a vanishing dipole. Therefore, in the calculation of a PMF between a pair of oppositely charged atoms that effect a non-

vanishing dipole in the simulation cell, the use of Ewald summation is likely to introduce artifacts.[131] Nevertheless, for the calculation of PMFs between charged species, both spherical truncation[132, 133] and Ewald summation [134, 135] techniques have been employed in the treatment of electrostatic interactions. In general, despite the problems associated with the use of the Ewald summation method to evaluate electrostatic interactions in the simulations involved in the determination of PMFs between pairs of charged species, where PMFs obtained using the spherical cut-off and Ewald techniques have been compared, those calculated via simulations with spherical truncation of electrostatic interactions have been shown to be adversely influenced by the treatment of electrostatic interactions,[136, 137] even resulting in physically unrealistic behaviour.

In this work, in order to reduce uncertainties, arising from the method of evaluating electrostatic interactions, in the accuracy of the PMFs determined, the electrostatic interactions were handled using a method developed by Wolf *et al*.,[74] who showed that artifacts in simulations involving spherical cut-offs are connected to the presence of net charges in the truncation spheres, as a consequence of the inclusion of a finite number of interactions. By achieving local neutralisation of charges, the Wolf method allows exact calculation of the coulombic interaction. Whilst the Wolf approach to the evaluation of electrostatic interactions has not previously been applied in the determination of a PMF between charged species, the suitability of the method as a means of calculating Coulombic interactions has been demonstrated for the simulation of ions in condensed systems.[138, 139]

3.2.3 PMF Calculation – Simulation Details

The PMFs for the approach of solute atoms in solvent were calculated from probability distribution functions obtained using Monte Carlo simulations in the canonical ensemble with the addition of artificial umbrella sampling potentials. Experimentally, Phase A of the cobalt succinates is synthesised at 348 K. Therefore, all umbrella sampling simulations involved in the calculation of PMFs intended for application in implicit-solvent simulations of the self-assembly of Phase A were

carried out at this temperature. The accuracy of PMFs determined using umbrella sampling and WHAM calculations is influenced by both the overlap of the windows in which biasing potentials are applied, and the sampling within the windows, as a consequence of the impact of such features of the simulated windows on the convergence of the PMF. The characteristics of the windows are controlled by interplay of the number of windows simulated, the force constants applied, and the number of steps simulated. In the simulation of a particular window, the magnitude of the force constant dictates the proportion of the configurations sampled which exist within the bounds of the window, and, consequently, the amount of sampling outside the window, which constitutes the overlap with adjacent windows exploited when optimally combining windows in the implementation of the WHAM algorithm. The number of steps simulated is crucial to the accuracy of the PMF since the range of the reaction coordinate examined in a given window must be adequately sampled so as to yield a probability distribution that is not dependent on the number of steps.

In other studies,[140-142] PMFs calculated between ions have been found to display minimal solvent effects and, accordingly, a relatively smooth profile beyond approximately 8.0 Å. In this work, separation distances were examined through the application of window sampling potentials starting between 1.0 Å and 2.0 Å, depending on the nature of the atom, and extending to 12.0 Å. This maximum separation distance was employed to make certain that all pertinent features were included in the PMFs. For all pairs of atoms, a window dimension of 0.1 Å was employed in the calculation of the PMF as a function of the distance separating the atoms. Following the suggestion by Wang *et al.*[143] that, for optimal application of the WHAM equations, approximately 30 % of the configurations sampled in each window should be inside the limits of the window, a force constant of 25 kcal mol^{-1} $Å^{-2}$ was used for the majority of umbrella sampling simulations. Where necessary, to ensure that the sampling was sufficiently confined to the locality of a particular window, higher force constants were utilised.

The use of narrow windows and the corresponding magnitude of the force constants applied were dictated by the requirement to minimise the number of steps required in

each umbrella sampling simulation; wider windows would necessitate the use of lower force constants and demand a greater number of steps to adequately sample the region of the reaction coordinate contained within the window. To ensure satisfactory sampling of the reaction coordinate, each window was simulated in the presence of the umbrella biasing potential for 300×10^6 steps, with distribution data collected every step. All simulations were carried out in a periodic cubic box with dimensions of 25 Å and containing explicitly represented water molecules in an amount corresponding to the experimental water density at the temperature studied. In the simulation of a particular window, solute atoms were initially placed in an equilibrated box of water such that the solutes were positioned centrally and separated by a distance equivalent to that at which the harmonic function used as the artificial biasing potential in the window is a minimum. The explicit water molecules in the simulation cell were represented using the TIP3P model,[94] which is suitable for use in conjunction with the utilised solute representations.[95] Use of the TIP3P model, which is a three-site representation of water, is associated with a lower computational cost than alternative water models involving higher numbers of interaction sites.

The potential parameters for the atoms in the solute ions and the solvent molecule were as shown in Table 3.1, and the cross-species interaction parameters were determined using the Lorentz-Berthelot mixing rules. The van der Waals interactions were expressed using the Lennard-Jones potential and, as previously described, electrostatic interactions were handled by means of the Wolf method. Both van der Waals and electrostatic interactions were cut-off at 12.0 Å. This distance corresponds to the maximum separation distance for which PMFs were calculated. Furthermore, since electrostatic interactions have been observed to be effectively short-ranged in condensed phase systems,[144] at this cut-off distance the interaction energies are expected to be negligible.

Table 3.1: Parameters for atoms in solute ions and solvent molecule. OW and HW represent, respectively, the oxygen and the hydrogen atoms in the water molecules.

Atom	σ (Å)	ε (K)	q (eu)	Ref.
Cobalt ion				
Co	2.186	14.392	+2.000	[93]
Succinate ion				
O	2.960	105.676	-0.800	[87]
C	3.750	52.838	+0.700	[87]
CH_2	3.905	59.380	-0.100	[87, 90]
TIP3P Water				
OW	3.151	76.526	-0.834	[94]
HW	0.000	0.000	+0.417	[94]

The umbrella sampling simulations were performed using a modified version of the multipurpose simulation code, Music,[96] whereas the WHAM analysis was performed using a program developed by the writer of this thesis. To further ensure the accuracy of the calculated PMFs, particular consideration was given to the convergence of the WHAM equations. Generally, solution of the WHAM equations is accepted as the point where the sum of the free energy constants is converged to a specified tolerance. However, this does not guarantee an equivalent convergence in the PMF because additional iteration of the coupled WHAM equations can effect changes of many kJ/mol in the PMF values. Thus, in this work, following the approach of Allen *et al.*,[145] full convergence of the unbiased PMF was judged to be attained only when every data point on the PMF satisfied the criterion of being converged to within 0.001 kcal/mol over 100 iterations. The WHAM calculations produce a smooth PMF profile with an uncertain baseline level. Thus, the PMFs obtained between pairs of atoms were shifted, in accordance with the practice adopted by Madhusoodanan and Tembe,[146] to coincide with the macroscopic coulombic potential at the maximum separation distance examined.

Errors in the PMFs were calculated using bootstrap analysis,[147] a resampling technique through which the reliability of a set of data can be evaluated. The use of

bootstrapping for the determination of errors associated with PMFs obtained using the WHAM algorithm has previously been implemented, and described in detail for calculation of PMFs obtained using the WHAM algorithm, by Hub *et al.*[148] Bootstrapping involves the generation of hypothetical probability distributions through random sampling of the data sets resulting from window simulations. The statistical uncertainty in the calculated PMF is determined through comparison of PMFs calculated from the hypothetical probability distributions. However, as a means of assessing the accuracy of the calculated PMFs, bootstrapping calculates only the statistical errors in the data sampled and cannot yield the extent of the uncertainty associated with the correlation between the specific region of the reaction coordinate that was intended to be sampled in a given window simulation and the actual location of the data represented in the sampled probability distribution.

3.3 Atom pair potentials of mean force

PMFs were calculated between all possible pairs of atoms belonging to the two solute ions using the method outlined previously. The positions of the atoms within the respective solute ions are illustrated in Figure 3.4. The PMFs obtained between the different pairs of atoms show distinct differences, which can be related to the parameters describing the atoms involved.

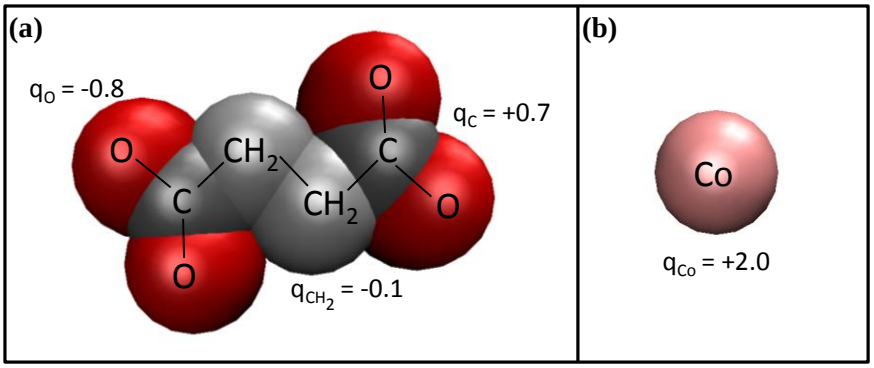

Figure 3.4: The atoms, of which PMFs are calculated between all possible pairs, are illustrated in relation to the solute species, where (a) shows the succinate ion, and (b) shows the cobalt ion.

As shown in Figure 3.5, the PMF between a pair of carbon atoms, as present in the succinate ion, is positive and therefore repulsive for all separation distances. This results from the repulsive nature of the coulombic attraction between the like-charged carbon atoms. The relatively featureless nature of the carbon-carbon PMF is a consequence of the hydrophobic nature of the carbon atom such that the water ordering has little impact on the relative positions of the carbon atoms. As demonstrated in Figure 3.5, the errors in the carbon-carbon PMF, calculated via bootstrap analysis method, are minimal to the extent that only the lines representing the upper and lower bounds of the error bars are visible. Since the errors in all of the PMFs calculated between pairs of atoms are of similar magnitude no error bars are displayed on subsequent representations of PMFs.

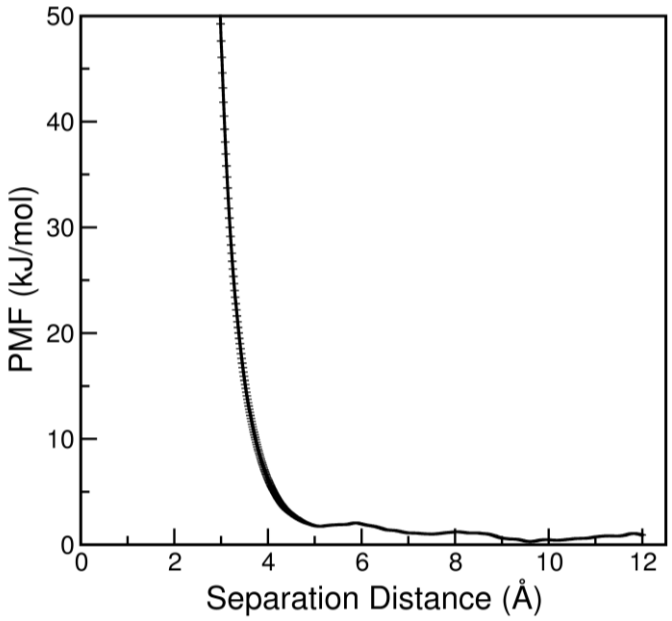

Figure 3.5: The potential of mean force between a pair of the carbon atoms in the succinate ion at 348 K, with error bars shown.

The PMF between the carbon atom and the methylene united atom in the succinate ion is shown in Figure 3.6. The distinct peaks and troughs apparent are characteristic of a PMF between a pair of oppositely charged species. The trough in the PMF at a separation distance of approximately 3.7 Å represents the contact distance (i.e., the distance at which the atoms are in direct contact with one another and, thus, no solvent exists between the atoms). The minimum in the PMF at approximately 6.6 Å

corresponds to the solvent-shared distance, at which the carbon and methylene species are separated by a single layer of solvent. The maximum between the contact distance and the solvent-shared distance at approximately 5.3 Å signifies the separation distance at the centre of the transition state between the two minima.[142]

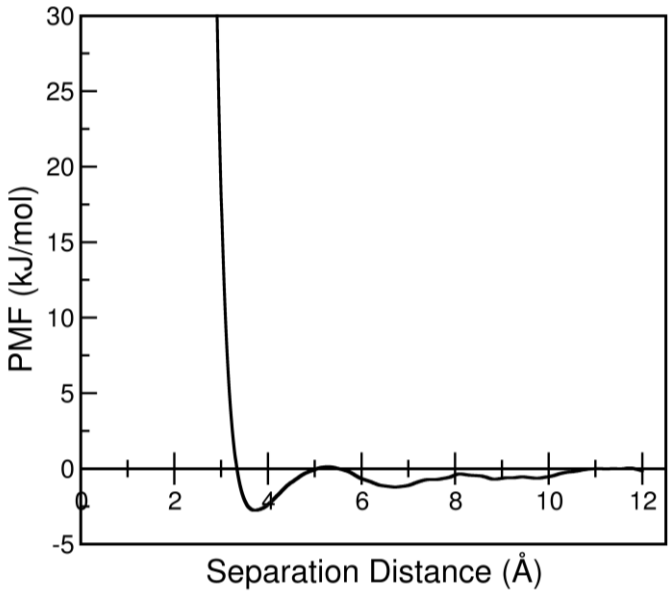

Figure 3.6: The potential of mean force between the carbon atom and the methylene united atom in the succinate ion at 348 K.

The PMF between the carbon atom in the succinate ion and the cobalt ion is shown in Figure 3.7. As with the carbon-carbon PMF shown in Figure 3.5, this PMF is positive for all values of separation distance, again reflecting the repulsive nature of the coulombic charge between the atoms. In comparison with the carbon-carbon PMF, the higher values of the carbon-cobalt PMF at shorter separation distances are a consequence of the greater charge on the cobalt ion. That the carbon-cobalt PMF is almost constant beyond approximately 7.5 Å demonstrates the negligible influence of the solvent ordering, at greater separation distances, on this pair of species.

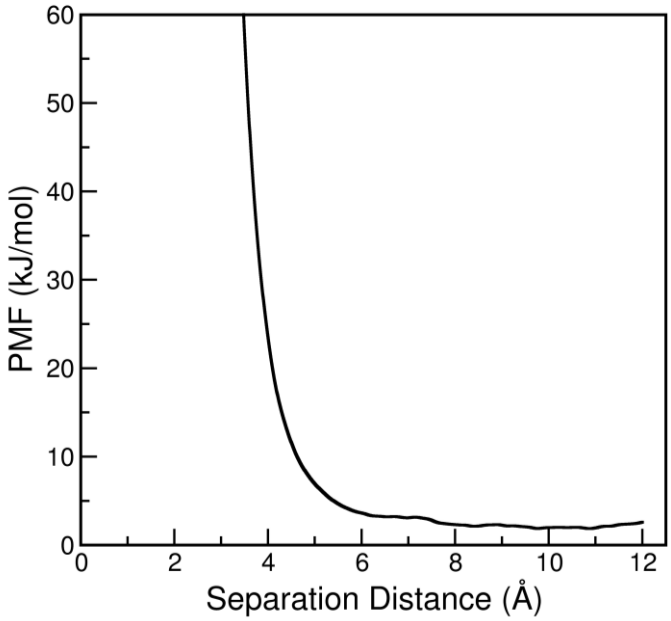

Figure 3.7: The potential of mean force between the carbon atom in the succinate ion and the cobalt ion at 348 K.

The PMF between the carbon atom in the succinate ion and the oxygen atom in the succinate ion is shown in Figure 3.8. As with the carbon-methylene PMF there is a clearly defined contact minimum, which, for the carbon-oxygen pair, occurs at a separation distance of approximately 3.1 Å. However, for the carbon-oxygen PMF, the solvent-shared minimum, at approximately 5.5 Å, is less distinct than that apparent for the C-CH_2 pair. That the solvent-shared minimum is less pronounced reflects both the hydrophobic nature of the carbon atom and the influence of the oxygen atom on the ordering of the solvent, whereby the water molecules arrange themselves with hydrogen atoms directed towards the oxygen solute atom. When the carbon and oxygen atoms are separated by distances greater than that at which there is a layer of solvent between the species, the oxygen atom will be surrounded by water molecules, from which the carbon atom is repelled. Therefore, beyond the distance corresponding to the second minimum in the C-O PMF, there are no configurations which are appreciably more or less favourable, with the effect that the second minimum is not especially well defined.

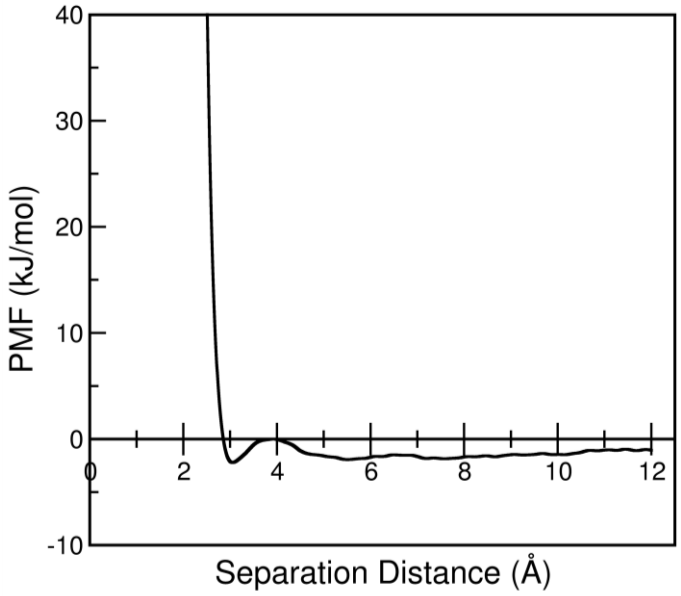

Figure 3.8: The potential of mean force between the carbon atom and the oxygen atom in the succinate ion at 348 K.

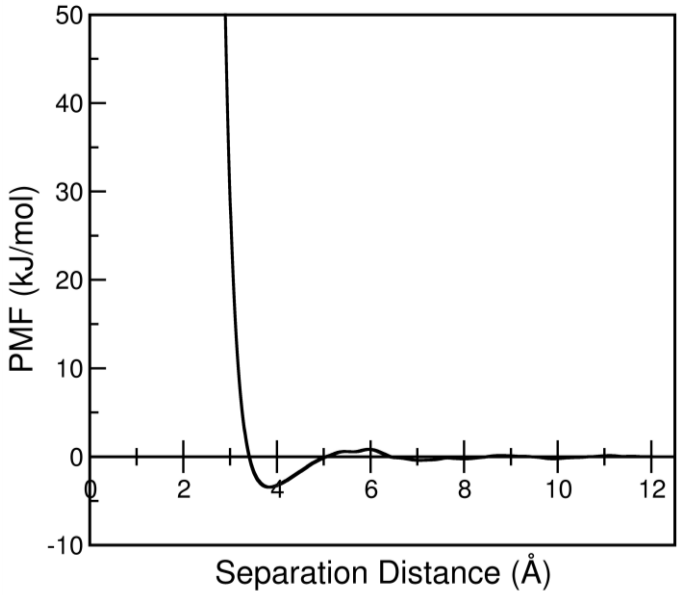

Figure 3.9: The potential of mean force between a pair of the methylene united atoms belonging to the succinate ion at 348 K.

The PMF for a pair of methylene united atom groups from the succinate ion is shown in Figure 3.9. In contrast with the other PMFs between like-charged atoms, the PMF for the CH_2-CH_2 united atom pair shows a distinct minimum at approximately 3.9 Å. This reflects the lower influence of the repulsive coulombic potential as a

consequence of the smaller magnitude of the charge on the methylene united atom. Therefore, the LJ interactions are more influential in the PMF calculation between the pair of methylene united atoms than for pairs of atoms with higher like-charges. The short-ranged nature of the CH_2-CH_2 PMF, such that there is little variation in the PMF beyond 6.5 Å, is also a consequence of the low charge on methylene because the electrostatic interactions between the methylene united-atoms are easily screened by the solvent molecules.

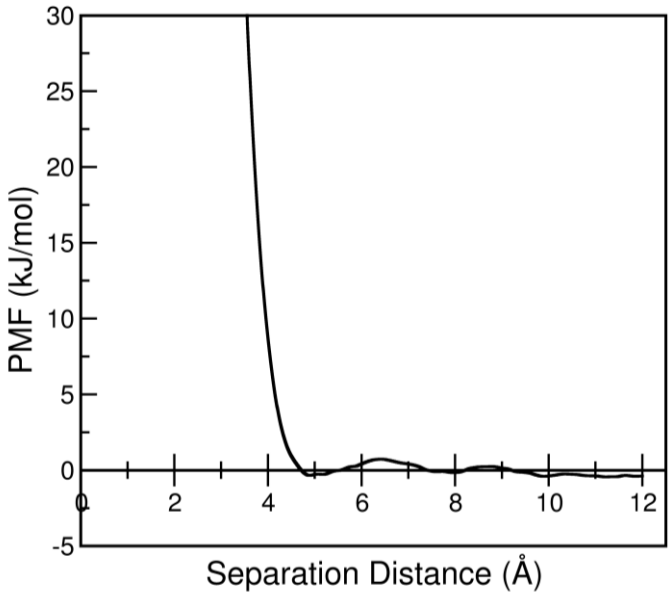

Figure 3.10: The potential of mean force between the methylene united atom in the succinate ion and the cobalt ion at 348 K.

The PMF between the methylene united atom in the succinate ion and the cobalt ion is shown in Figure 3.10, which demonstrates that there is minimal variation in the PMF with separation distance for the CH_2-Co pair. This results from the fact that the species are relatively weakly interacting. Nevertheless, since CH_2 and Co are oppositely-charged, the coulombic interaction between the species is attractive in nature. Therefore, beyond the combined LJ sigma parameter value ($\sigma_{CH_2\text{-}Co}$ = 3.046 Å), both the coulombic and non-coulombic interactions effect an attraction between CH_2 and Co. Despite this, the PMF between CH_2 and Co shows that the species repel each other for separation distances less than approximately 4.7 Å, a distance which exceeds even that at which the LJ minimum occurs (3.418 Å). This is likely

to be a consequence of the proclivity of the cobalt to form an octahedral complex with six coordinated water molecules because the work required to supplant these solvent molecules exceeds that which can be gained by the methylene united atom and the cobalt adopting configurations in which the species are more closely positioned.[149]

The PMF between the methylene united atom in the succinate ion and the oxygen atom in the succinate ion is shown in Figure 3.11. Despite the relatively strong attractive (beyond 3.43 Å) non-coulombic interaction between methylene and oxygen, the CH_2-O PMF shows that the repulsive electrostatic interaction arising from the like-charges of the species is dominant, and produces a gradual diminishment in the PMF values between approximately 3.3 Å and 6 Å.

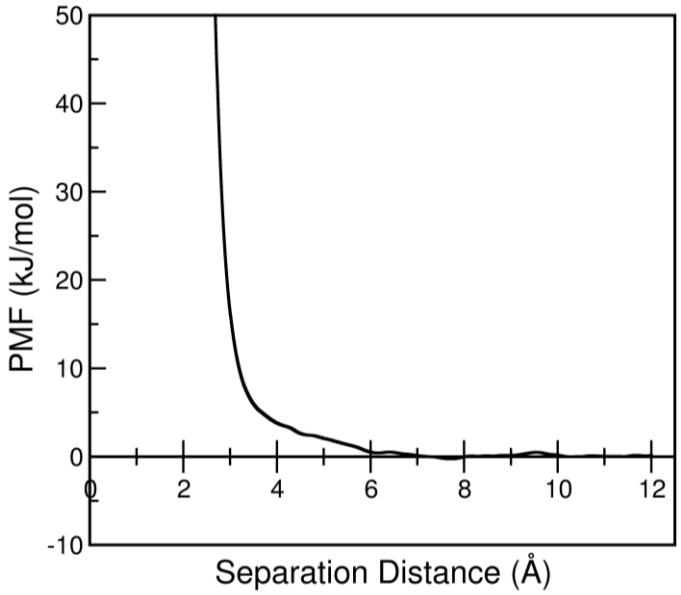

Figure 3.11: The potential of mean force between the methylene united atom and the oxygen atom in the succinate ion at 348 K.

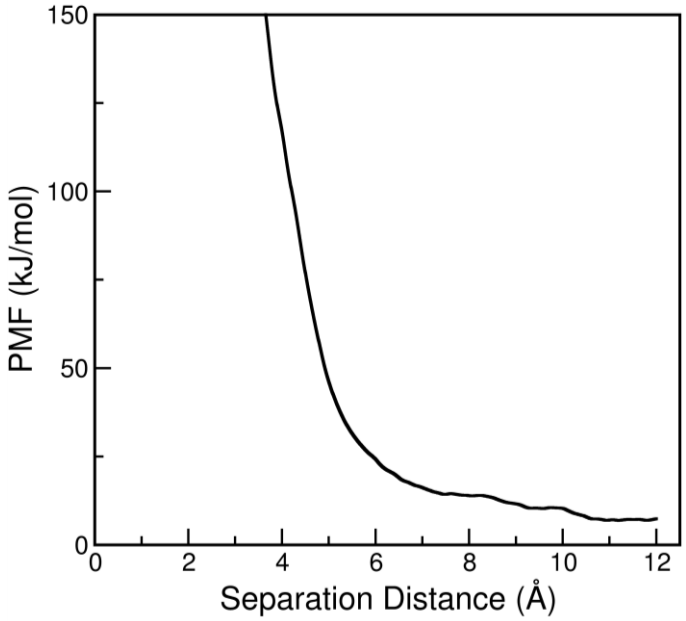

Figure 3.12: The potential of mean force between two cobalt ions at 348 K.

The PMF between two cobalt ions is shown in Figure 3.12. The characteristics of the Co-Co PMF are primarily related to the tendency for the cobalt ion to surround itself with six water molecules in an octahedral configuration. Configurations representative of those sampled in four different window simulations along the reaction coordinate are shown in Figure 3.13. The arrangement of the cobalt pair and coordinating water molecules shown in Figure 3.13a was sampled in the simulation of the window in which the biasing potential applied was used to constrain the sampling close to 2.0 Å. Because of the proximity of the cobalt ions, in order that both ions have a full complement of coordinating water molecules, two water molecules must be coordinated in cooperation between the ions. The obligation to share two water molecules presents great difficulties to the cobalt ions with respect to the preservation of the octahedral arrangement of water molecules. Therefore, this configuration is extremely unfavourable and, as a consequence, is a very high-energy state, as shown in Figure 3.12.

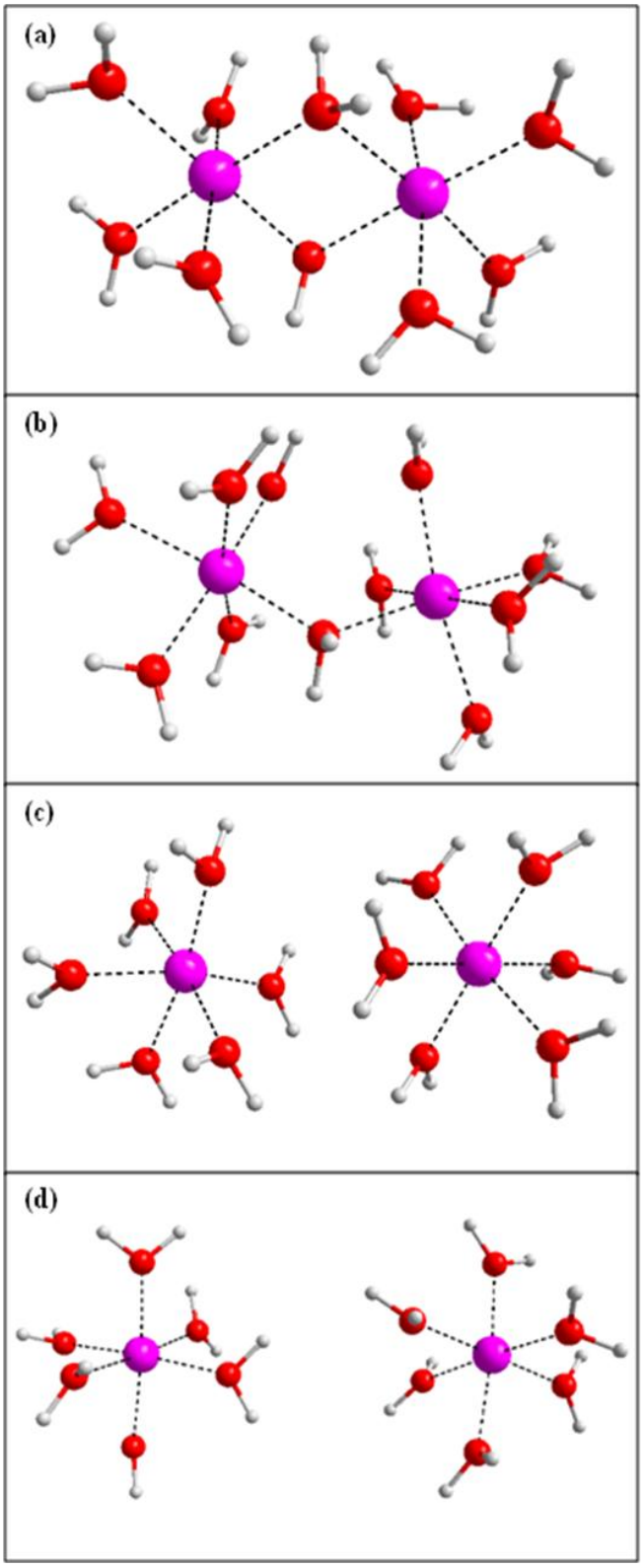

Figure 3.13: Configurations representative of those sampled when a biasing potential is applied to constrain the Co-Co pair at: (a) 2.0 Å; (b) 3.0 Å; (c) 5.0 Å; (d) 7.0 Å. Spheres are not representative of atom size. (Co, pink; O, red; H, white). Dashed lines show the arrangement of the water molecules coordinating the cobalt ions.

The configuration shown in Figure 3.13b is typical of those sampled when a biasing potential constraining the pair of cobalt ions at a separation of 3.0 Å is applied. With a greater distance separating the cobalt ions, only one water molecule must be shared between the cobalt ions to satisfy the requirement that the cobalt ions are both coordinated to six water molecules. One mutual water molecule represents a lesser problem for the maintenance of the octahedral arrangement of the water molecules coordinating the cobalt ions but is nevertheless impedimental, with the result that, at a separation distance of 3.0 Å, the configurations sampled are high-energy states. Typical of the configurations sampled when a pair of water-coordinated cobalt ions is constrained at a separation distance of 5.0 Å, the arrangement shown in Figure 3.13c demonstrates that, at this separation distance, the cobalt ions are sufficiently far apart that both ions can coordinate six respective water molecules. Consequently, as shown in Figure 3.12, at a separation distance of 5.0 Å, the configurations sampled are much more favourable and represent much lower-energy states. However, the proximity of the cobalt ions is such that there remains some hindrance to the formation of perfectly octahedral arrangements of cobalt and water because the water molecules coordinated to one cobalt ion will impact upon the arrangement of the water molecules surrounding the other cobalt ion. With greater distance between the cobalt ions, the configuration shown in Figure 3.13d, which is characteristic of those sampled when the cobalt-cobalt distance was constrained at 7.0 Å, demonstrates that the water molecules surrounding one cobalt ion have less bearing on the arrangement of the water molecules coordinated to the opposing cobalt in the pair. Therefore, when separated by a distance of 7.0 Å, the cobalt ions are able to adopt much more favourable configurations and, correspondingly, represent comparatively low-energy states, as shown in Figure 3.12.

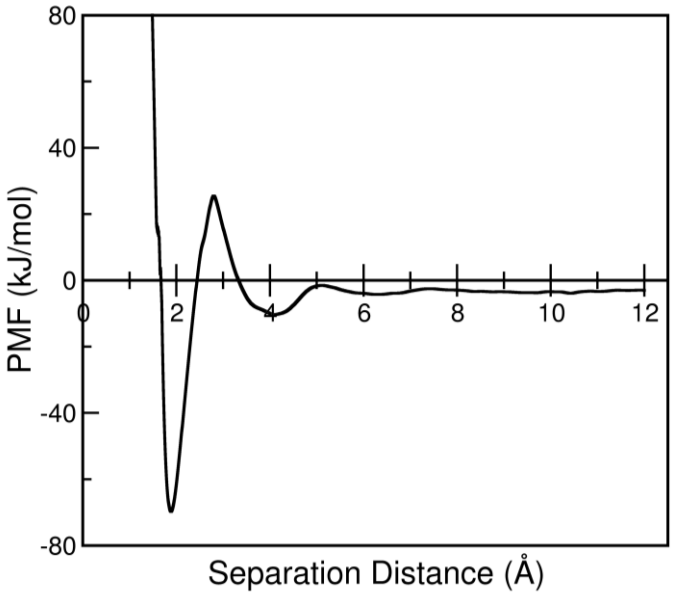

Figure 3.14: The potential of mean force between the cobalt ion and the oxygen atom in the succinate ion at 348 K.

The PMF between the cobalt ion and the oxygen atom in the succinate ion is shown in Figure 3.14. The PMF between Co and O is influenced both by the tendency for Co to coordinate water molecules in an octahedral configuration and by the ability of the oxygen atom to influence the ordering of the solvent molecules. The deep minimum, at approximately 1.9 Å, in the Co-O PMF indicates the contact distance for Co and O. The depth of the minimum reflects that the 'contact' configuration is highly favourable, as a consequence of the large opposing charges of the species, which are, therefore, strongly attracted. The configuration shown in Figure 3.15a, is representative of those sampled in the simulation of a window biased to sample separation distances close to 1.9 Å. This configuration demonstrates that, at the contact-separation distance, the solute oxygen atom constitutes one of the six oxygen atoms coordinated to the cobalt in order to fulfil the requirement for an octahedral configuration, where, otherwise, oxygen atoms are constituents of the solvent water molecules.

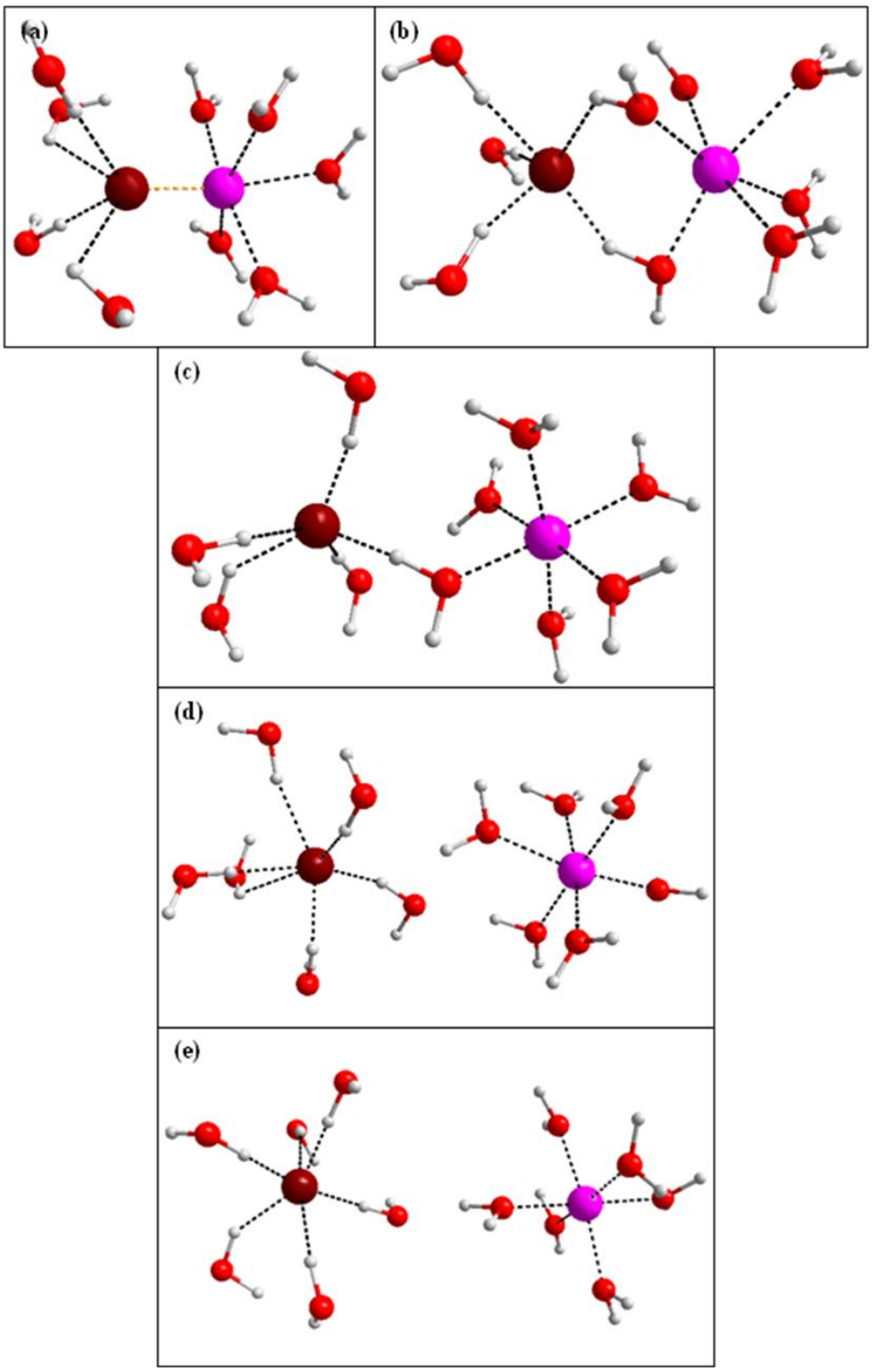

Figure 3.15: Configurations representative of those sampled when a biasing potential is applied to constrain the Co-O pair at: (a) 1.9 Å; (b) 2.8 Å; (c) 4.1 Å; (d) 5.2 Å; (e) 6.3 Å. Spheres are not representative of atom size. (Co, pink; O (solute), dark red; O (solvent), red; H, white). Black dashed lines show the arrangement of the solvent water molecules surrounding the Co and O solutes and the orange dashed line represents that the solute O is coordinated to Co.

The peak, at approximately 2.8 Å, in the PMF between Co and O represents the transition between the contact distance and the solvent-shared minimum, which is apparent at approximately 4.1 Å. The configuration shown in Figure 3.15b is typical of those sampled when cobalt and oxygen are constrained at a separation distance of 2.8 Å. At this separation distance the gap between the cobalt and oxygen species is such that two of the water molecules coordinated to the cobalt are associated in the ordering of the water molecules surrounding the oxygen atom. Because of the combined influence, on the nearby solvent molecules, of the cobalt and oxygen at such proximity, this configuration is particularly unfavourable and represents a very high-energy state in the Co-O PMF. Figure 3.15c exemplifies a configuration representative of those sampled when the cobalt-oxygen pair is constrained at a separation of 4.1 Å. In this arrangement only one of the water molecules coordinated to the cobalt is directly involved in the ordering of the first layer of solvent surrounding the oxygen atom. Therefore, at a separation distance of 4.1 Å, the cobalt and oxygen are sufficiently far apart that the water molecule in question can adopt a position, with respect to both the cobalt and oxygen species and the other solvent molecules, whereby the configuration represents a relatively low-energy state, as shown in Figure 3.14. The minimum in the Co-O PMF at a separation distance of approximately 6.3 Å represents the solvent-separated minimum and signifies the distance at which the cobalt and oxygen species are both surrounded by an individual, intact layer of solvent. As shown in the configuration displayed in Figure 3.15e, which is characteristic of the arrangements sampled in the simulation of the window in which a harmonic function is applied to constrain the distance between Co and O close to 6.3 Å, both cobalt and oxygen are surrounded by water molecules. At this separation distance, the water molecules constituting the shells of solvent surrounding Co and O can position themselves so as to have little impact on one another and the resulting configuration is relatively low-energy. In contrast, the configuration typical of those sampled in a window simulation involving the application of a constraint distance of 5.2 Å, as shown in Figure 3.15d, demonstrates that, at a shorter separation distance, the water molecules involved in the layer of solvent surrounding one solute species have an influence on the water molecules contained in the shell of solvent encasing the other solute. Therefore, as shown in

Figure 3.14, configurations in which cobalt and oxygen are separated by 5.2 Å are less favourable and represent higher-energy states than those in which the Co-O separation distance is 6.3 Å.

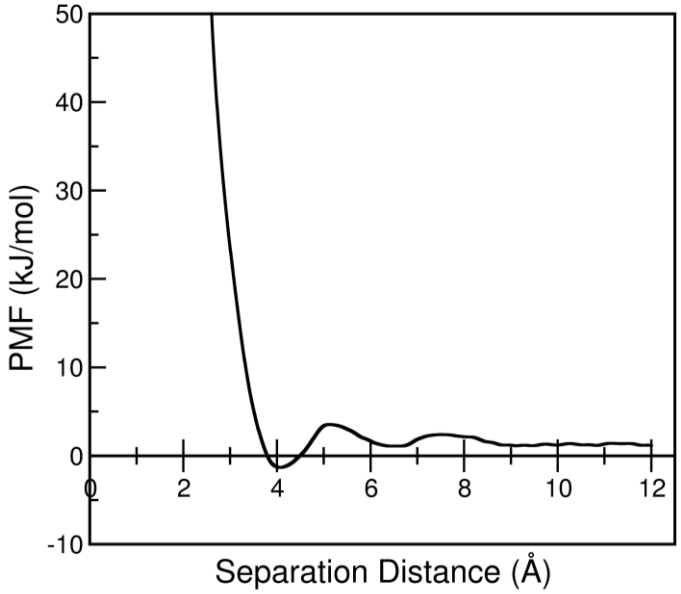

Figure 3.16: The potential of mean force between a pair of the oxygen atoms belonging to the succinate ion at 348 K.

The PMF between a pair of the oxygen atoms from the succinate ion is shown in Figure 3.16. The O-O PMF is largely influenced by the effect of the oxygen atom on the ordering of the water molecules. As illustrated in Figure 3.16, the oxygen-oxygen PMF shows a minimum at a separation distance of approximately 4.0 Å. A representative configuration of the pair of oxygen atoms constrained at a separation distance of 4.0 Å is shown in Figure 3.17a. This arrangement demonstrates that, at such an oxygen-oxygen separation, of the water molecules involved in the arrangement of the layers of solvent surrounding the oxygen atoms, two water molecules are positioned such that they are constituents of the solvation shells of both oxygen atoms. These two water molecules effectively bridge, and therefore stabilise, the oxygen pair to form a favourable configuration.

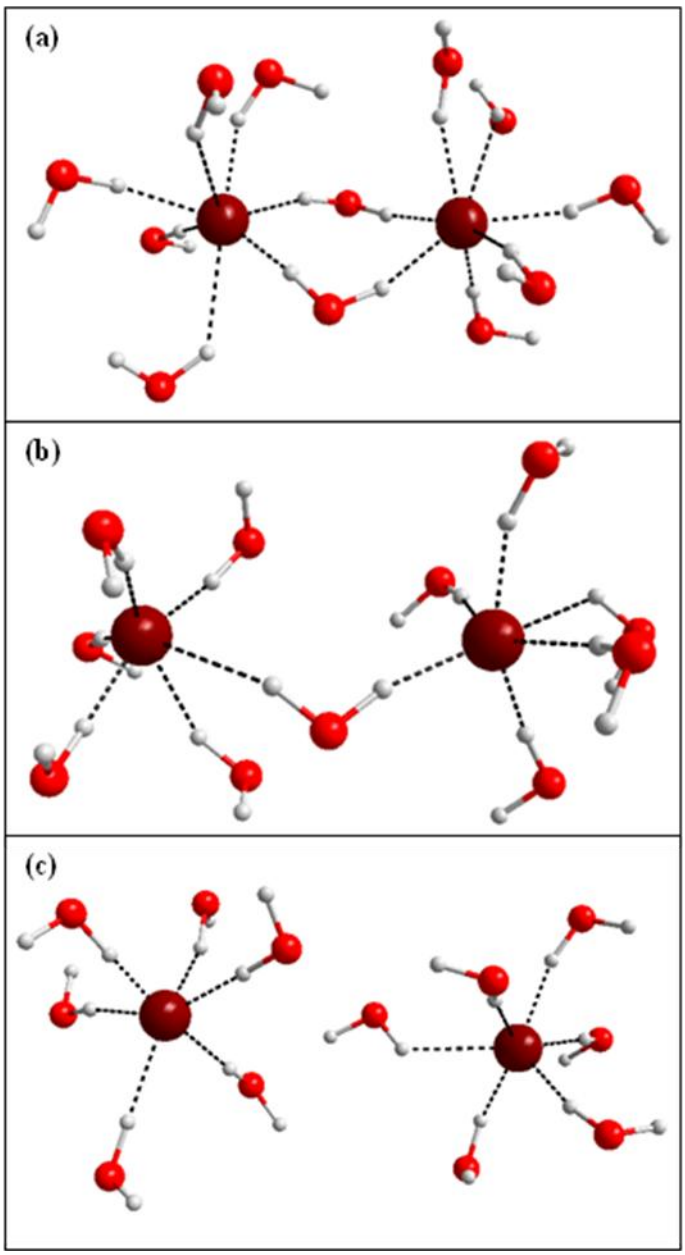

Figure 3.17: Configurations representative of those sampled when a biasing potential is applied to constrain the O-O pair at: (a) 4.0 Å; (b) 5.2 Å; (c) 6.5 Å. Spheres are not representative of atom size. (O (solute), dark red; O (solvent), red; H, white). Dashed lines show the arrangement of the water molecules surrounding the oxygen solute atoms.

Beyond the minimum in the oxygen-oxygen PMF, all separation distances are associated with positive PMF values. This reflects the repulsion between the like-charged oxygen atoms. Within the repulsive region of the O-O PMF, the peak at a separation distance of approximately 5.2 Å indicates that when the oxygen atoms are separated by this distance, the resulting configurations are particularly unfavourable.

As demonstrated in the configuration shown in Figure 3.17b, which is typical of those sampled when the oxygen atoms are constrained at a distance of 5.2 Å, at this separation distance only one solvent molecule is involved in the primary solvation shells of both oxygen atoms. However, at an oxygen-oxygen separation distance of 5.2 Å, for a water molecule to be jointly associated with the solvation shells of both oxygen atoms, the water molecule must be positioned a non-optimal distance from the oxygen atoms, with the result that the configuration is a relatively high-energy state. When separated by a greater distance, the oxygen atoms have individual solvation shells that do not impact on each other, as illustrated in Figure 3.17c, which shows a configuration representative of those sampled when the distance between the oxygen atoms was constrained at 6.5 Å. Thus, configurations involving a pair of oxygen atoms separated by a distance of 6.5 Å are more favourable and represent a lower-energy state, with the effect that there is a minimum in the O-O PMF at 6.5 Å, corresponding to the solvent-separated minimum.

3.4 Validation of methods implemented

That all of the PMFs calculated between pairs of the atoms involved in the succinate and cobalt ions have relatively smooth profiles and tend to constant values at distances less than the maximum distance at which the PMF was determined indicates that the outlined method for obtaining PMFs between pairs of ions yields accurate results. Nevertheless, for further validation of the implementation of the methods, the influence of the cut-off distance was examined through comparison with a PMF calculated using a longer cut-off distance. Since the features of the Co-O PMF were most distinct, the effect of the implementation of a longer cut-off distance was assessed for the cobalt-oxygen pair. Effective analysis of the influence of the cut-off distance required that the influence of an appreciably greater cut-off distance was examined. This necessitated the use of a larger unit cell so that the cut-off distance was not greater than half of the length of the unit cell. Therefore, the Co-O PMF calculated in a cubic unit cell with dimensions of 25 Å and using a 12.0 Å cut-off distance was compared with the equivalent PMF determined using simulations performed in a cubic unit cell with dimensions of 30 Å, with interactions

cut-off at 14.0 Å. As shown in Figure 3.18, for completeness, the aforementioned PMFs were assessed alongside the cobalt-oxygen PMF obtained in simulations carried out in a cubic unit cell with dimensions of 30 Å, where the evaluation of interaction potentials involved the application of a 12.0 Å cut-off distance.

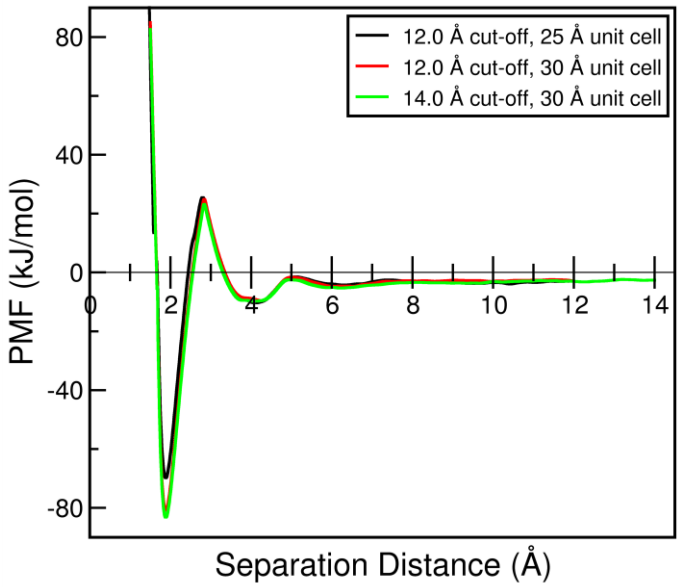

Figure 3.18: Evaluation of the influence of cut-off distance on the Co-O potential of mean force at 348 K.

As shown in Figure 3.18, at longer separation distances, the Co-O PMF calculated using a 14.0 Å cut-off distance is virtually identical to that obtained using the shorter cut-off. Therefore, the use of a 12.0 Å cut-off distance for the evaluation of interaction potentials is adequate for the calculation of PMFs. The differences in the PMF values at the Co-O contact distance at a separation distance of 1.9 Å are a consequence of the favourableness of this configuration. Due to the propensity of the cobalt-oxygen pair to exist in the contact configuration, there is some difficulty associated with accurately sampling the less favourable configurations that involve the cobalt-oxygen pair separated by distances marginally shorter or longer than the contact distance. Therefore, there is a greater degree of uncertainty associated with the PMF values at distances close to the contact distance than the PMF values at other regions of the reaction coordinate. However, the extent of the disparity, shown in Figure 3.18, in the different PMF values at the contact distance is not of sufficient

magnitude as to be significant in comparison with the actual value of the PMF at this distance.

In addition, in all simulations involved in the calculation of the PMFs between the pairs of the constituents of the succinate and cobalt ions, the solvent molecules were represented by the TIP3P model for water. This water model was employed in the derivation of the potential parameters for the cobalt ion.[93] However, the parameters for the atoms involved in the succinate ion[87, 90] were obtained in conjunction with the TIP4P[94] representation of the water molecule. Nevertheless, since the two different water models generate results of sufficient resemblance as to be considered virtually transposable,[95] the use of the TIP3P model was thought to represent an acceptable approximation. In order to confirm this, the O-O PMF obtained using the TIP3P water model was compared with the equivalent PMF determined with the solvent represented using the TIP4P description, as shown in Figure 3.19. This comparison was effected for the oxygen-oxygen pair since, of the atoms in the succinate ion, the oxygen atom is the most influenced by the solvent.

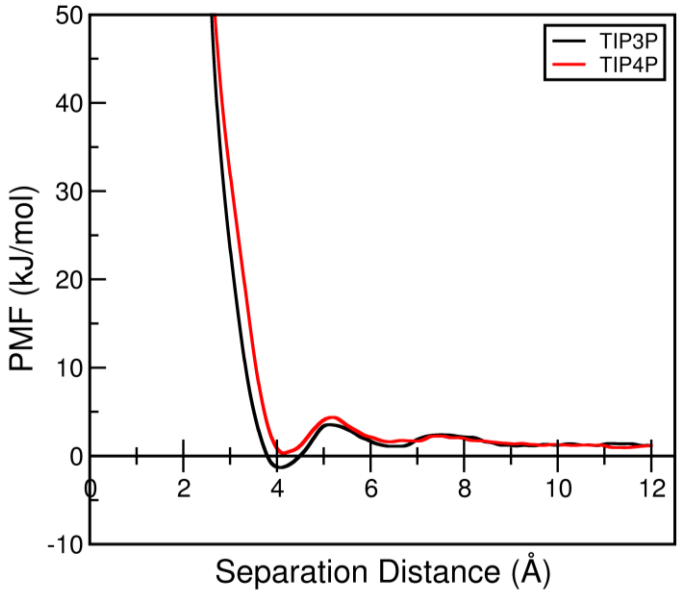

Figure 3.19: Comparison of the oxygen-oxygen PMFs calculated using, respectively, the TIP3P and TIP4P water models.

As shown in Figure 3.19, with water molecules described using the TIP4P model (the parameters for which are shown in Table 3.2), the solvent-separated minimum is slightly less favourable than the corresponding configuration involving solvent molecules represented using the TIP3P parameters. Thus, the use of the four-site TIP4P water model produces only a small difference, at short separation distances, in the O-O PMF, as compared with the use of the TIP3P representation. Furthermore, the locations of the maxima and minima are not influenced by the solvent representation method, indicating that the use of the TIP3P model is a reasonable approximation.

Table 3.2: TIP4P water parameters.[94] M represents a dummy atom in the water model.

Atom	σ (Å)	ε (K)	q (eu)
TIP4P Water			
OW	3.154	78.009	0.000
HW	0.000	0.000	+0.520
M	0.000	0.000	-1.040

As described in Section 3.2.2, there is much question relating to the appropriateness of the usage of particular methods of evaluating electrostatic interactions with respect to the calculation of PMFs between species of ionic character. The effects of the different approaches to the calculation of electrostatic interactions are examined in Section 3.5.

3.5 Consideration of electrostatic interactions treatment

In studying different treatments of electrostatic interaction in relation to the calculation of PMFs, a comparison was effected of the cobalt-oxygen PMFs obtained using a number of techniques for the representation of electrostatic interactions. Thus, as shown in Figure 3.20, the Co-O PMF calculated from simulations involving electrostatic interactions handled using the Wolf method was compared with the corresponding PMFs determined in simulations in which electrostatic interactions were evaluated using Ewald summation and a spherical truncation method. The

spherical truncation method applied involved truncating the interaction and shifting to zero at the cut-off distance. At low separation distances, all of the evaluated PMFs are concurrent on the locations of the maxima and minima, with only small differences in the PMF values at the distances at which these features are apparent. At large separation distances, there are only marginal differences between the PMFs obtained using the Wolf method and those calculated with electrostatic interactions handled using the Ewald technique. However, the PMF obtained using spherical truncation shows significant undulations and does not converge to a constant value at large separation distances. Therefore, with respect to the calculation of a PMF between charged species, Figure 3.20 indicates that the Wolf and Ewald approaches give very similar results, which differ from the results of the spherical truncation method. Whilst this comparison cannot provide a definite indication of the accuracy or otherwise of the Wolf approach employed in this work, the fact that there is need for careful consideration of the method of evaluating electrostatic interactions is clear.

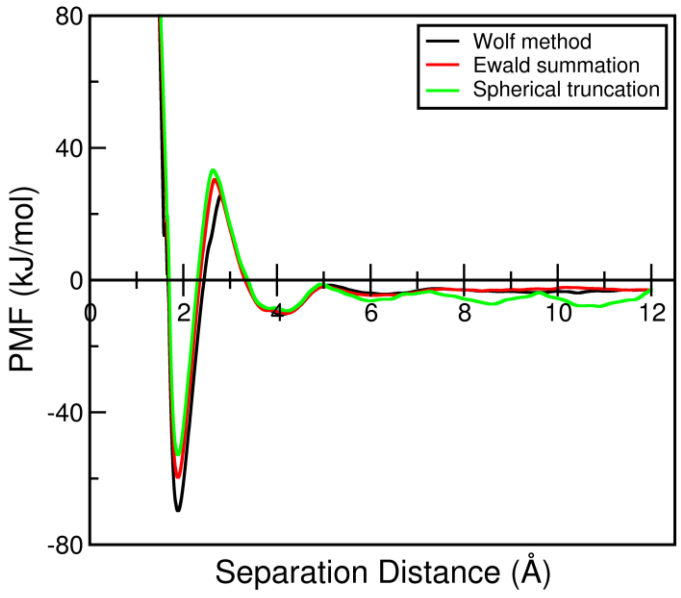

Figure 3.20: Evaluation of the influence, on the cobalt-oxygen potential of mean force at 348 K, of different methods of calculating electrostatic interactions.

3.6 Summary

To enable the development of a computationally practicable means of simulating the synthesis of MOFs an implicit-solvent technique has been formulated. A PMF approach was employed to allow MOF-synthesis simulations to be carried out without the explicit representation of solvent molecules. In utilising the cobalt succinate materials as a test system such that the implicit-solvent method was required to represent the interactions between succinate and cobalt ions, PMFs were calculated between all pairs of the constituent atoms. These PMFs are presented and the forms of the profiles are discussed. All of the calculated PMFs are reinforced by physically reasonable explanations for the features and characteristics displayed, indicating that the outlined method can be employed to obtain accurate PMFs between species of ionic nature. The smoothness of the PMF profiles indicates that all regions of the reaction coordinate were sufficiently well sampled and demonstrates that the bin size employed was appropriate to ensure convergence. The adequacy of the sampling of the reaction coordinates in the calculation of PMFs was further exhibited by the near negligibility of the errors in the PMFs, which indicates that the errors in the probability distributions sampled are minimal.

The suitability of the method used to calculate the PMFs between the charged species was additionally demonstrated through investigation of the influence of certain parameters. In conjunction with the illustration of the negligible impact of the application, with respect to the calculation of interaction potentials, of a longer cut-off distance, the fact that the PMFs determined tend to constant values at higher separation distances shows that the distance at which interaction potentials were cut-off was suitable. Comparison of PMFs determined using TIP3P and TIP4P water representations indicates that the explicit solvent molecules can be reasonably described using the TIP3P model for the pairs of ionic species.

CHAPTER 4

Evaluation of the potential-of-mean-force approach

The successful implementation of an implicit-solvent method requires that the technique can reasonably accurately replicate the explicit-solvent behaviour. In Chapter 3, details are given of a potential-of-mean-force approach to the implicit representation of the solvent in a system that undergoes self-assembly. Before the approach described can be applied in MOF-synthesis simulations involving an implicit representation of the solvent, the capabilities of the method must be assessed.

In implicit-solvent simulations involving PMFs calculated between all possible pairs of the constituents of the solutes there are a number of potential sources from which might arise inconsistencies with the corresponding explicit-solvent simulations. PMFs calculated between individual atoms represent the implicit-solvent potential

between the atoms in infinite dilution and, therefore, do not take into account the other atoms that exist within the solute species. Thus, there is a possibility that, in simulations of only the fully solvated constituent-pair, as involved in the calculation of a PMF, the favourabilities of respective configurations sampled might not, as a consequence of steric hindrance or bonding interactions, correspond to those that result when the corresponding atoms exist within the molecular environment. In addition, accurate determination of the interactions between the solute species by pairwise summation of the PMFs is dependent on the suitability of the approximation that, in the calculation of interactions, the constituent PMFs are additive. A further lack of consistency is that in implicit-solvent simulations the solute species are present at higher concentrations than in the simulations from which PMF data were obtained. Therefore, the infinite-dilution, pairwise PMFs might not adequately reflect the interactions. Thus, the approach of determining the effective interactions between the atom constituents of the solute species does not necessarily result in potentials with the capability to reproduce the interactions and the associated behaviour observed in an explicit-solvent simulation. There is merit, therefore, in evaluating, in relation to the implicit-solvent method developed in this work, any impact such possible inconsistencies might have. This can be achieved through comparison of the functioning of the implicit-solvent method with the performance of the employed explicit-solvent approach, which, as described in Chapter 2, is based on the use of well-founded molecular potentials.

4.1 Evaluation of explicit-solvent and implicit-solvent representations

To establish the suitability of the obtained PMFs for representing the potentials between the initial reactants in the cobalt succinate system (i.e. the succinate ion and the cobalt ion), explicit-solvent and implicit-solvent representations were compared for each respective pair of the ions to determine whether representative configurational sampling is reproduced using pairwise PMFs. However, the incentive for the development of an implicit-solvent method – that the explicit-solvent simulations are unfeasibly long – also impinges on this comparison, since

thorough evaluation of the configurations sampled with the solvent modelled explicitly would require extensive simulation. Therefore, in order to obtain the correct configurational sampling in the explicit-solvent simulations within a reasonable time-scale, constraining, harmonic potentials were imposed to restrict the sampling close to a specified separation distance. Thus, for explicit-solvent and implicit-solvent simulations, the application of equivalent harmonic potentials between pairs of ions allowed comparison, of the distributions of the distances separating the ions, and enabled evaluation of any differences, arising from the method of solvent representation, in the configurations sampled.

For both explicit-solvent and implicit-solvent cases, data relating to the configurations sampled were collected using MC simulations in the canonical ensemble,[56] which is described in Chapter 2, with the positions of the species involved in the simulation altering by displacement and, where appropriate, rotational moves. In the implicit-solvent sampling simulations, interactions between the ions were determined by pairwise summation of the PMFs between the constituent atoms. For consistency, the methods employed in the explicit-solvent sampling simulations corresponded to those implemented in the simulations employed in the determination of the PMFs and, thus, van der Waals interactions between species were represented using the LJ potential and electrostatic interactions were treated using the Wolf[74] method. The initial configuration for the explicit-solvent simulation was generated by immersing the pair of ions of interest in a box of equilibrated water molecules with a density corresponding to the experimental water density at the temperature at which both the implicit-solvent and explicit-solvent sampling simulations were carried out. The temperature corresponded to that at which the PMFs utilised in the implicit-solvent simulation were obtained (348 K).

The ions were initially positioned centrally in the simulation box so as to be separated by a distance equivalent to the minimum in the constraining potential. Thus, the starting configurations for equivalent explicit-solvent and implicit-solvent sampling simulations contained the ions in identical relative positions. This is

illustrated in Figure 4.1, which also demonstrates the action of the constraining, harmonic potential between the solutes.

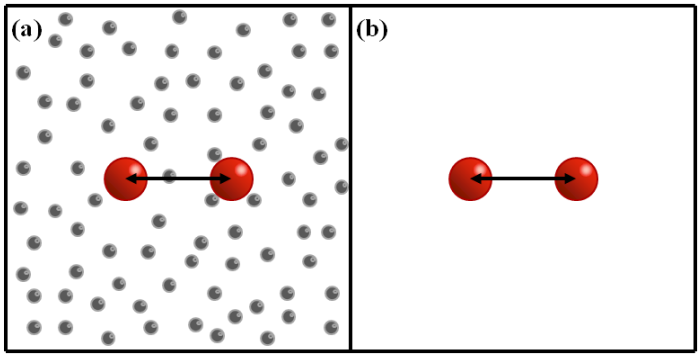

Figure 4.1: Representative initial configurations in (a) explicit-solvent, and (b) implicit-solvent sampling simulations. In these images, the red spheres represent the solute species, the grey spheres represent the solvent molecules, and the double-headed arrow reflects the action of the constraining potential.

The magnitude of the constraining potential, as imposed by the value of the force constant in the harmonic function constraining the ions, dictates the extent to which the ions can separate, and influences the number of simulation steps required to obtain the final distribution of separation distances between the ions in the configurations sampled. Therefore, so as to avoid the need to undertake extremely lengthy explicit-solvent sampling simulations, a series of sampling simulations involving decreasing values of force constant were performed. This ensured determination of the lowest force constant for which the explicit-solvent sampling simulations could be carried out within a reasonable time-frame. The use of this lowest practicable force constant allowed the comparison of the explicit-solvent and implicit-solvent methods by means of distributions of separation distance resulting from sampling simulations least influenced by constraining potentials. These distributions of the separation distances between specified locations on both members of a given pair of ions were determined by accumulating data from each step of a simulation in which possible configurations were sampled. The probability density distributions employed as a means of comparing the behaviour of the ion pairs in explicit-solvent and implicit-solvent representations were determined by examining the convergence of the distributions produced as the sampling simulation progressed. For both implicit-solvent and explicit-solvent sampling simulations the

final distributions were taken as those that were unchanged by the inclusion of more steps. The explicit-solvent and implicit-solvent configurational sampling of all possible pairs of the solute ions involved in the cobalt succinate system is evaluated in Section 4.2.

4.2 Configurational sampling of pairs of solute ions

4.2.1 Cobalt-cobalt configurational sampling

The comparison of the explicit-solvent and implicit-solvent representations for the cobalt-cobalt ion pair is effectively an assessment of the accuracy of the Co-Co PMF. To examine the range of separation distances included in the PMF, for the cobalt-cobalt pair, the explicit-solvent and implicit-solvent representations were compared by determining distributions of the distances separating two Co^{2+} ions constrained so as to be located at separation distances close to 3, 5, 7 and 9 Å.

Comparisons of the probability density distributions resulting from explicit-solvent and implicit-solvent sampling simulations involving cobalt ions constrained at a separation distance of 3 Å using force constants of 10, 5 and 1 kcal·mol^{-1}·Å^{-2} are shown in Figure 4.2, and the Co-Co PMF is given in Figure 4.3. With a force constant of 10 kcal·mol^{-1}·Å^{-2}, in the explicit-solvent sampling simulation, configurations are sampled in which the cobalt ions are separated by distances between 3.8 and 4.6 Å, and for force constants of 5 and 1 kcal·mol^{-1}·Å^{-2}, the separation distances between the cobalt ions in the sampled configurations range, respectively, from 4.1 to 5.2 Å and from 4.8 to 6.5 Å. The differences in the ranges of separation distances sampled are a consequence of both the steep negative gradient and the strongly repulsive nature of the Co-Co PMF at low separation distances, such that the application of a lower force constant in the harmonic function constraining the cobalt ions has the effect of allowing the ions to occupy a wider range of separation distances located further from the constraint distance. Therefore, the distributions resulting from the applications of the three different force constants

are necessary for evaluation of the ability of the Co-Co PMF to replicate the cobalt-cobalt behaviour when the solvent is represented explicitly.

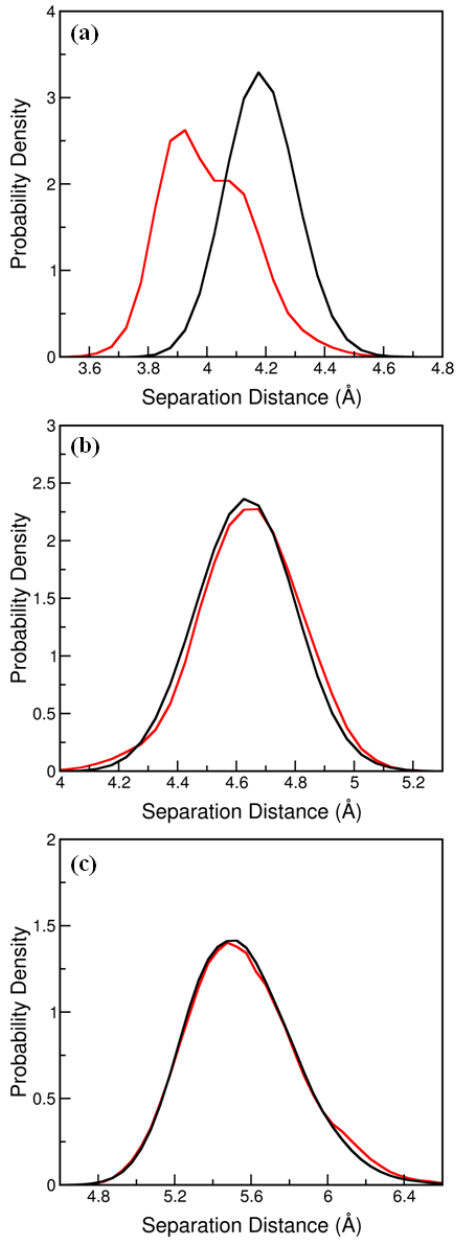

Figure 4.2: Comparison of implicit-solvent and explicit-solvent probability density distributions of the sampled separation distances between two cobalt ions constrained at a distance of 3 Å using force constants of (a) 10 kcal·mol^{-1}·Å^{-2}, (b) 5 kcal·mol^{-1}·Å^{-2}, and (c) 1 kcal·mol^{-1}·Å^{-2}. In these images the implicit-solvent distributions are represented by red lines and the explicit-solvent distributions are represented by black lines. The implicit-solvent distributions incorporate data from 1×10^{6} sampling steps and the explicit-solvent distributions incorporate data sampled from (a) 8×10^{8} steps, (b) 4×10^{9} steps, and (c) 2×10^{9} steps.

As shown in Figure 4.2a, when a force constant of 10 kcal·mol^{-1}·Å^{-2} is applied the distributions of separation distances sampled in the implicit-solvent and explicit-solvent simulations differ significantly. This indicates that the Co-Co PMF does not accurately replicate the explicit-solvent behaviour of a pair of cobalt ions for the entire range of separation distances sampled when this force constant is imposed. However, the agreement, shown in Figures 4.2b and 4.2c, between the explicit-solvent and implicit-solvent probability density distributions obtained when force constants of 5 and 1 kcal·mol^{-1}·Å^{-2} are employed demonstrates that the Co-Co PMF can effectively represent the explicit-solvent interactions in the corresponding ranges of separation distances.

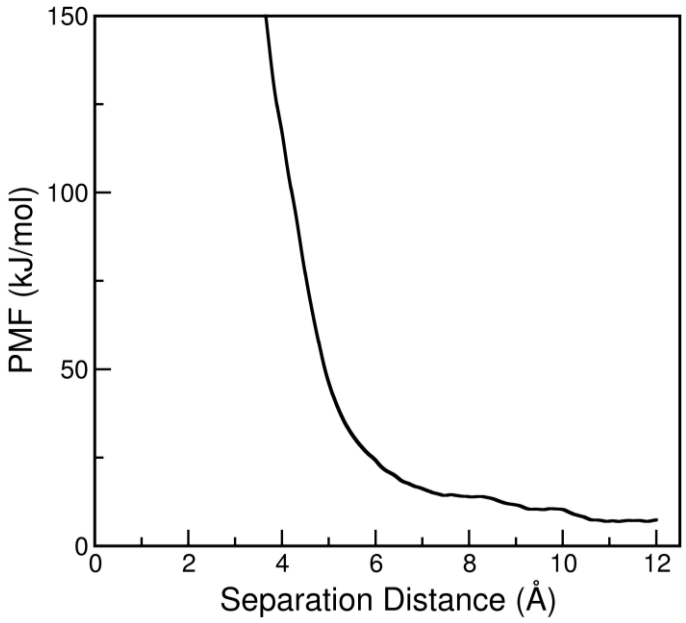

Figure 4.3: The potential of mean force between two cobalt ions at 348 K.

The comparison of the explicit-solvent and implicit-solvent probability density distributions of separation distances between a pair of cobalt ions constrained, using a force constant of 1 kcal·mol^{-1}·Å^{-2}, to sample distances close to 5 Å is shown in Figure 4.4. In this case, the cobalt-cobalt ion pair exists in configurations in which the separation distance between the cobalt ions ranges from 5.0 to 7.6 Å and there is reasonable agreement between the distributions obtained for explicit and implicit representations of solvent.

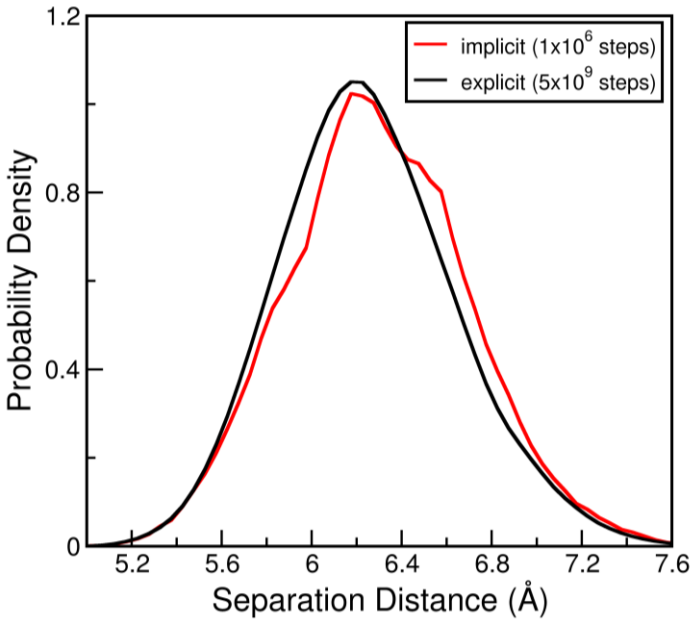

Figure 4.4: Comparison of implicit-solvent and explicit-solvent probability density distributions of the sampled separation distances between two cobalt ions constrained at a distance of 5 Å using a force constant of 1 kcal·mol^{-1}·Å^{-2}.

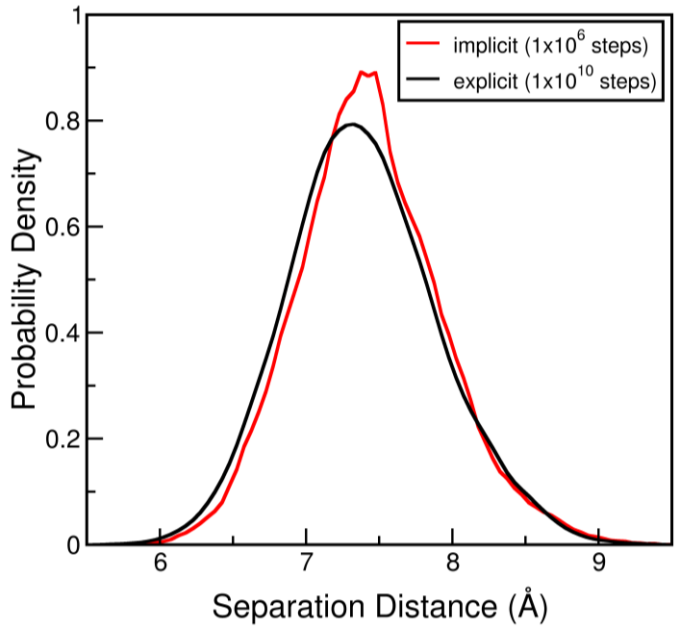

Figure 4.5: Comparison of implicit-solvent and explicit-solvent probability density distributions of the sampled separation distances between two cobalt ions constrained at a distance of 7 Å using a force constant of 1 kcal·mol^{-1}·Å^{-2}.

As illustrated in Figure 4.5, when the cobalt ions are constrained, through the application of a force constant of 1 kcal·mol^{-1}·Å^{-2}, at a separation distance of 7 Å, the probability density distributions of separation distances sampled in simulations with

explicit-solvent and implicit-solvent representations show fairly good concordance. The implementation of the aforementioned constraint leads to the sampling of configurations in which the separation distances between the cobalt ions range from 5.8 to 9.0 Å. In order to study the ability of the Co-Co PMF to represent the behaviour, in explicit-solvent, of a pair of cobalt ions constrained at a separation distance of 9 Å, a force constant of 2 kcal·mol^{-1}·Å^{-2} was required to ensure that, in the configurations sampled, the ions were separated by distances shorter than that at which the interactions were cut-off. As shown in Figure 4.6, with the cobalt ions constrained as described, the sampling simulations involved configurations in which the ions were separated by distances ranging from 7.8 to 10.5 Å. The resulting probability density distributions from explicit-solvent and implicit-solvent simulations agree fairly well for the range of separation distances involved.

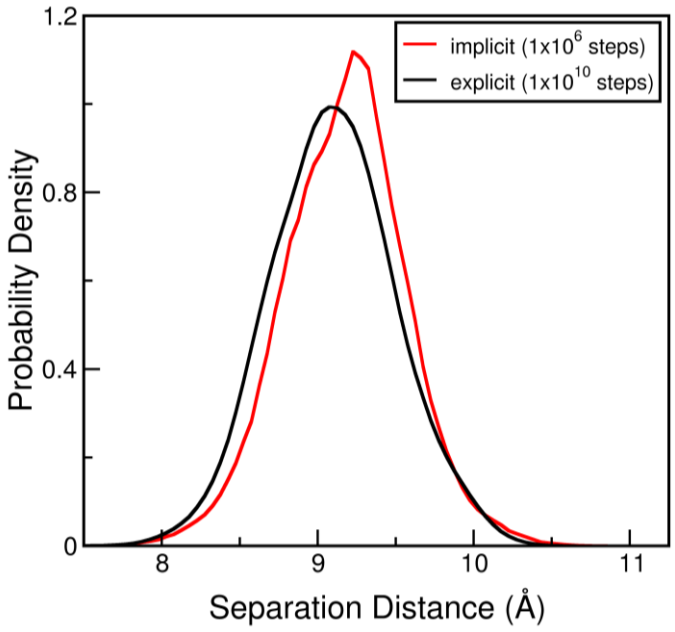

Figure 4.6: Comparison of implicit-solvent and explicit-solvent probability density distributions of the sampled separation distances between two cobalt ions constrained at a distance of 9 Å using a force constant of 2 kcal·mol^{-1}·Å^{-2}.

In summary, with the exception of the range of separation distances sampled when a force constant of 10 kcal·mol^{-1}·Å^{-2} was used to constrain a pair of cobalt ions close to 3 Å, the Co-Co PMF can represent, with reasonable accuracy, the explicit-solvent behaviour of the cobalt-cobalt pair. Therefore, disparity between the explicit-solvent

and implicit-solvent representations was only significant at separation distances where the interaction between the cobalt ions is highly repulsive, and the associated configurations are extremely unfavourable. Consequently, the differences between the probability density distributions resulting from explicit-solvent and implicit-solvent sampling simulations will be of little consequence in self-assembly simulations with the cobalt-cobalt interactions represented using the PMF approach. The agreement between the distributions obtained from explicit-solvent and implicit-solvent simulations for the ranges of separation distances sampled in all other constrained simulations indicates that the cobalt-cobalt interaction can be represented using the Co-Co PMF and provides further evidence of the accuracy of the PMF.

For the comparisons, shown in Figure 4.2 and Figures 4.4 – 4.6, of the probability density distributions obtained from explicit-solvent and implicit-solvent sampling simulations, the final distributions of the separation distances between the cobalt ions in the configurations sampled were obtained, in a matter of minutes, within 1×10^{6} steps when the solvent was represented implicitly. However, with the solvent was modelled explicitly the number of steps required to determine the final distribution of separation distances was typically of the order of 10^{9}-10^{10}, and the sampling simulations took several weeks to complete. This both shows the power of the PMF approach and confirms the impossibility of studying the self-assembly of MOFs using explicit-solvent simulations.

4.2.2 Cobalt-succinate configurational sampling

For thorough analysis of the capability of the implicit-solvent method in relation to the interaction between the succinate and cobalt ions, the behaviour of the ion pair was evaluated separately for the cobalt ion constrained parallel to the carbon chain in the succinate ion and for the cobalt ion constrained perpendicular to the chain of carbon atoms in the succinate ion. The arrangements of the cobalt and succinate ions in these individual assessments are shown in Figure 4.7. The respective comparisons of the probability density distributions for explicit-solvent and implicit-solvent representations are detailed in Sections 4.2.2.1 and 4.2.2.2.

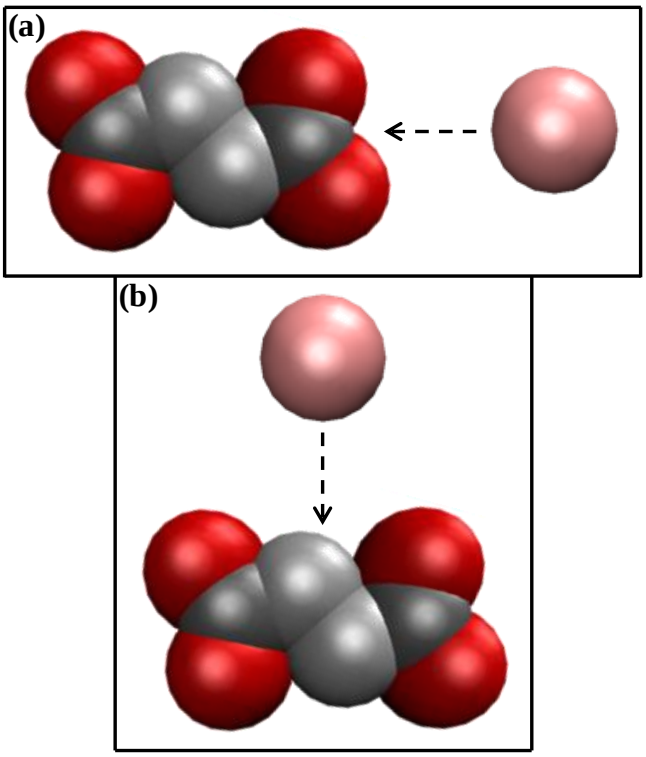

Figure 4.7: The cobalt-succinate ion pair was evaluated for the cobalt ion approaching the succinate ion (a) parallel to the carbon chain, and (b) perpendicular to the chain of carbon atoms.

4.2.2.1 Cobalt ion constrained parallel to succinate ion carbon chain

With the cobalt ion constrained parallel to the chain of carbon atoms in the succinate ion, the assessment of the ability of the implicit-solvent method to replicate the explicit-solvent behaviour was intended primarily as a judgement of the accuracy of the Co-O PMF. Because the terminal carbon in the succinate ion is attached to two identical oxygen atoms, which both interact with the cobalt ion, the separation distance data from which the probability density distributions were constructed was collected between the cobalt ion and both oxygen atoms. Therefore, to effectively compare the explicit-solvent and implicit-solvent methods, and ensure that in the configurations sampled the cobalt ion interacted principally with the oxygen atoms, constraints of equal magnitude were applied between the cobalt ion and both oxygen atoms in the succinate ion, as shown in Figure 4.8.

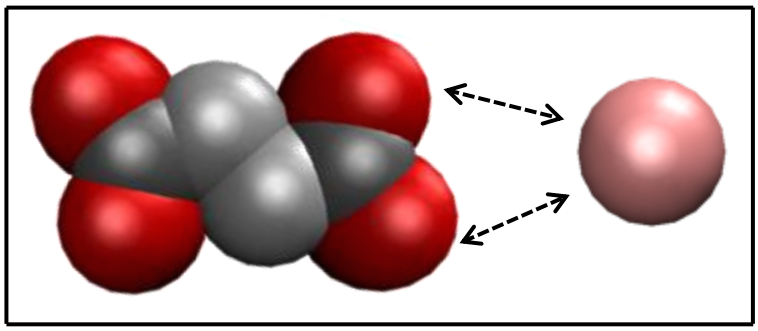

Figure 4.8: The cobalt ion was constrained parallel to the carbon chain in the succinate ion by means of constraints (represented by dashed arrows) between the cobalt ion and both oxygen atoms at one end of the succinate ion.

With the cobalt ion constrained at a distance of 3 Å, using a force constant of 2 $kcal \cdot mol^{-1} \cdot Å^{-2}$, from both of the oxygen atoms at one terminus of the succinate ion, the sampling simulations yielded the probability density distributions of separation distance shown, for explicit-solvent and implicit-solvent representations, in Figure 4.9. This comparison clearly demonstrates that there are significant differences between the configurations sampled when the solvent is modelled explicitly and those sampled when PMFs between the atom constituents of the ions are employed to represent the solvent implicitly. In the explicit-solvent distribution shown in Figure 4.9, the peak of high magnitude evident between approximately 1.8 and 2.1 Å indicates that configurations involving the cobalt ion positioned a distance within this range from one of the oxygen atoms in the succinate ion are particularly favourable. In such configurations, the distance between the cobalt ion and the other oxygen atom at the same terminus of the succinate ion ranges from approximately 3.1 to 4.2 Å, as demonstrated by the second peak in the explicit-solvent distribution. The greater width of the peak at higher separation distances reflects that, while there is a distinct optimum distance at which the cobalt ion exists from an oxygen atom, there are a range of positions the cobalt ion can occupy at the favourable proximity. That the implicit-solvent method does not replicate the explicit-solvent behaviour, particularly at short separation distances, demonstrates that there is a problem associated with the use of the pairwise PMFs to represent the interactions between the succinate ion and the cobalt ion.

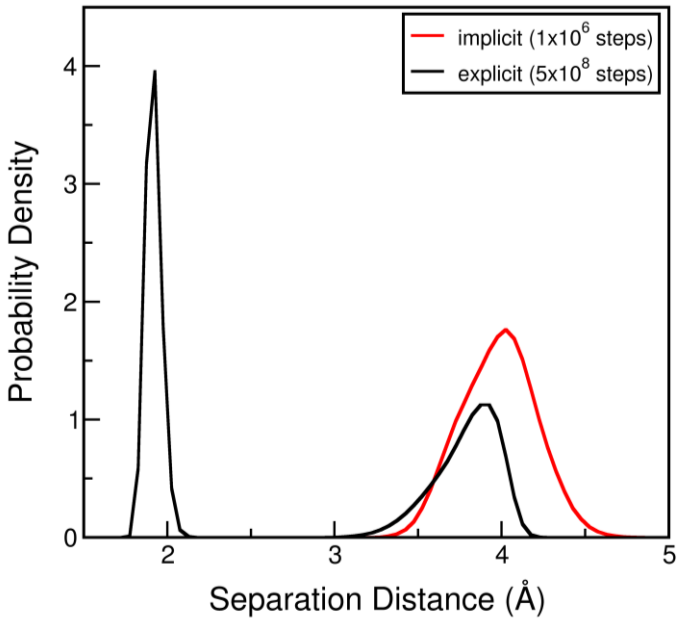

Figure 4.9: Comparison of implicit-solvent and explicit-solvent probability density distributions of the separation distances sampled in simulations involving the succinate-cobalt ion pair constrained between the cobalt ion and the oxygen atoms at one terminus of the succinate ion, with constraint distances of 3 Å and force constants of 2 kcal·mol^{-1}·Å^{-2}.

The reason for the failure of the implicit-solvent sampling simulations to yield the peak corresponding to the optimum separation distance between the cobalt ion and an oxygen atom in the succinate ion was investigated by examining the locations occupied, through the course of sampling simulations, by the cobalt ion in relation to the succinate ion. With the succinate ion at a fixed site in the centre of the simulation box and the cobalt ion free to move in the absence of any constraints, the positions sampled by the cobalt ion were recorded for explicit-solvent and implicit-solvent simulations. The distributions of the point locations occupied by the cobalt ion are shown in Figures 4.10a and 4.10b, respectively, for explicit-solvent and implicit-solvent representations.

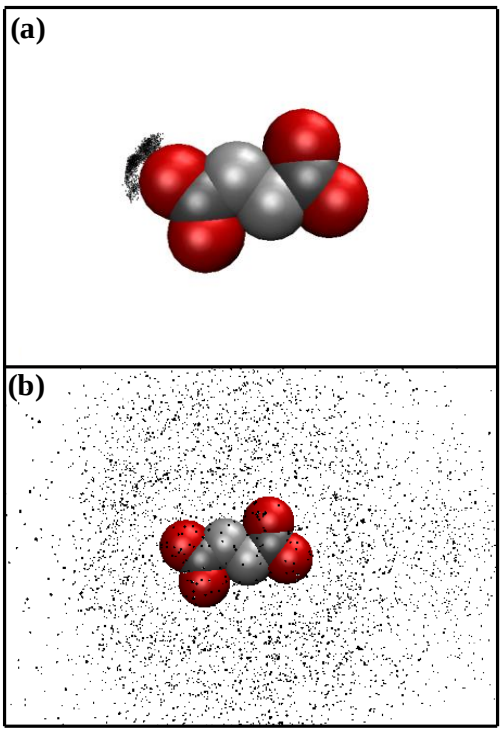

Figure 4.10: Distributions of the positions occupied by the cobalt ion in relation to a fixed-position succinate ion for: (a) explicit-solvent simulations, and (b) implicit-solvent simulations, where black points represent the locations sampled by the cobalt ion.

With the solvent modelled implicitly, cobalt-ion position data were collected during a simulation involving 1×10^{6} steps and with an explicit-solvent representation data were obtained for a simulation lasting 2×10^{9} steps. As shown in Figure 4.10a, with an explicit-solvent representation, the positions occupied by the cobalt ion in relation to the fixed-position succinate ion are distributed in the vicinity of only one of the oxygen atoms in the succinate ion. This is a result of the explicit representation of the solvent molecules, which surround the succinate ion and inhibit the acceptance of sampling moves that would allow the cobalt ion to be positioned close to the other oxygen atoms in the succinate ion. Therefore, the distribution resulting from the unconstrained explicit-solvent simulation by no means represents of all of the sites that the cobalt ion might occupy. However, the distribution of locations sampled by the cobalt ion in the implicit-solvent simulation is a good exemplification of all the locations in which the cobalt ion might exist when the interactions between the cobalt ion and the succinate ion are represented using PMFs between the constituents of the ions. Therefore, since the implicit-solvent distribution of cobalt locations does

not include any points closer than 3.225 Å from any of the oxygen atoms in the succinate ion, the fact that all of the points which constitute the explicit-solvent distribution are within this distance of their nearest oxygen atom indicates that some feature of the PMFs prevents the cobalt ion from closely approaching the succinate ion, even when no constraining potential is applied.

Thus, the absence, in the implicit-solvent distribution, of a peak corresponding to the optimum separation distance for the cobalt ion in relation to the succinate ion results from the cobalt ion being prevented from occupying regions close to the oxygen atoms in the succinate ion. Situated at the termini of the succinate ion, the oxygen atoms in the ion exist in environments not greatly dissimilar from the isolated atom which was involved in the calculation of PMFs. Therefore, any restriction to the approach of the succinate ion by the cobalt ion was unlikely to result from characteristics of the PMF between the cobalt ion and the oxygen atom. However, being attached to three other species, the carbon atom in the succinate ion exists in a significantly different environment to that employed in the determination of the C-Co PMF. Thus, since the PMF between the carbon atom and the cobalt ion does not account for the effect of the atoms attached to the carbon atom, the possibility for inconsistency when the C-Co PMF is implemented with respect to the entire succinate ion is not insignificant.

For the implicit-solvent method to accurately replicate the explicit-solvent behaviour of the cobalt-succinate ion pair an acceptable solution to the problem associated with the utilisation of a solely PMF-based approach must be devised. Since the situation of the carbon atom amongst the other species in the succinate ion is such that the impact of the C-Co interaction on the behaviour of the cobalt-succinate ion pair is minimal, the PMF between the carbon atom and the cobalt ion was replaced with a hard-sphere potential. As shown in Equation 4.1, with the interaction between the carbon atom and the cobalt ion represented by means of a hard-sphere potential, U_{HS}, the potential at a separation distance, r_{ij}, below or equal to a specified cut-off distance, r_{cut}, is of infinite magnitude and for separation distances greater than the cut-off distance the potential is zero.

$$U_{HS}(r_{ij}) = \begin{cases} \infty, & r_{ij} \leq r_{cut} \\ 0, & r_{ij} > r_{cut} \end{cases} \quad \textbf{4.1}$$

The form of the hard-sphere potential used to represent the C-Co interaction is shown alongside the C-Co PMF in Figure 4.11. The hard-sphere potential in this comparison results from the application of a cut-off distance equivalent to the cross-species Lennard-Jones sigma value for the carbon-cobalt pair. As demonstrated in Figure 4.11, the hard-sphere potential between the carbon atom and the cobalt ion has lower magnitude than the equivalent PMF for all separation distances beyond the hard-sphere cut-off. Thus, the hard-sphere potential presents less obstruction to the cobalt ion approaching the oxygen atoms in the succinate ion.

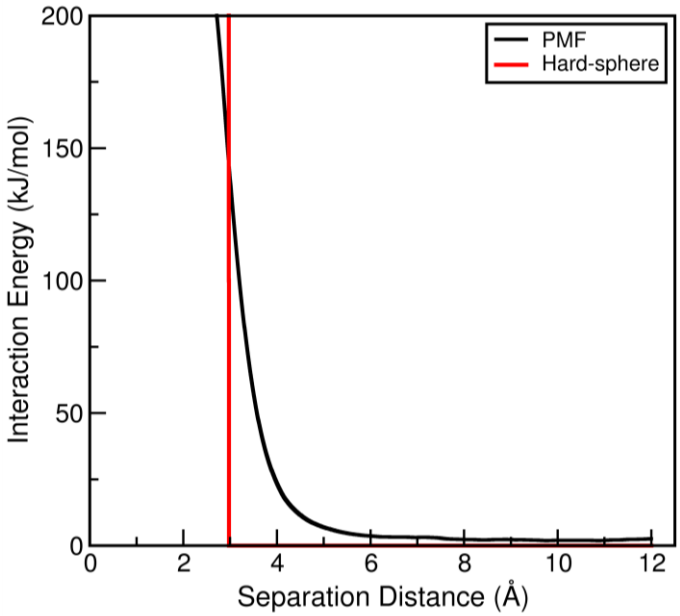

Figure 4.11: Comparison of the PMF and the hard-sphere potential for the carbon-cobalt interaction.

The suitability of the use of a hard-sphere potential to represent the interaction between the carbon atoms in the succinate ion and the cobalt ion was assessed by determining the positions, surrounding the succinate ion, available to the cobalt ion. The distribution of the locations occupied by the cobalt ion with respect to the

succinate ion is shown in Figure 4.12. In determining this distribution, with the exception of the carbon-cobalt potential, which was represented by the hard-sphere potential, the interactions were represented by means of the relevant PMFs. As in the determination of similar distributions, the cobalt ion and succinate ion were unconstrained. As shown in Figure 4.12, the positions, in relation to the oxygen atoms in the succinate ion, of the sites occupied by the cobalt ion when the carbon-cobalt interaction is represented by means of a hard-sphere potential correspond to the relative locations of the ions in the configurations sampled in the equivalent explicit-solvent simulation.

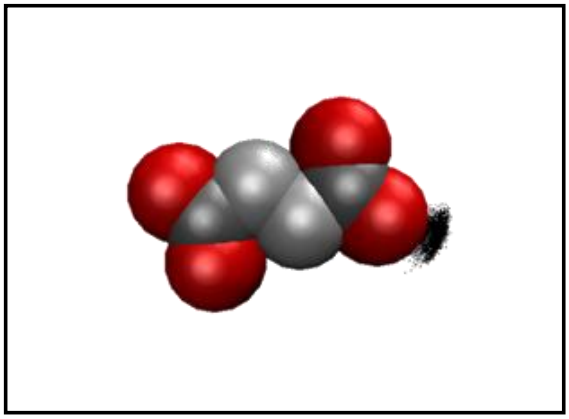

Figure 4.12: Distribution of the positions occupied by the cobalt ion in relation to a fixed-position succinate ion, where black points represent the locations sampled by the cobalt ion. The solvent was represented implicitly and a hard-sphere potential was employed in place of the C-Co PMF.

For a clearer comparison of the suitability of employing a hard-sphere potential in place of the C-Co PMF in the implicit-solvent sampling simulations, probability density distributions resulting from implicit-solvent and explicit-solvent representations were evaluated from sampling simulations involving the cobalt-succinate ion pair constrained, as previously described, using force constants of 2 $kcal \cdot mol^{-1} \cdot Å^{-2}$ and constraint distances of 3 Å. The resulting distributions of the separation distances between the cobalt ion and the constrained oxygen atoms in the succinate ion in the configurations sampled are shown in Figure 4.13. The presence of the peak spanning separation distances of approximately 1.8 – 2.1 Å in the implicit-solvent probability density distribution demonstrates that by representing the carbon-cobalt interaction by means of a hard-sphere potential there is no hindrance to

the cobalt ion closely approaching the oxygen atoms in the succinate ion. That the height and baseline width of this peak coincide well with the equivalent peak in the explicit-solvent distribution indicates the better capacity of the hard-sphere potential, in comparison with the C-Co PMF, to represent the carbon-cobalt interaction.

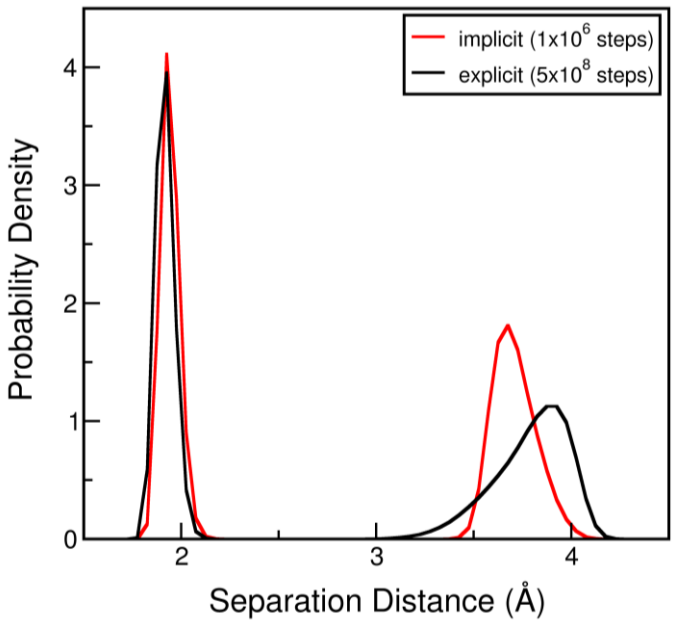

Figure 4.13: Comparison of implicit-solvent and explicit-solvent probability density distributions of the separation distances sampled in simulations involving the succinate-cobalt ion pair constrained between the cobalt ion and the oxygen atoms at one terminus of the succinate ion, with constraint distances of 3 Å and force constants of 2 kcal·mol^{-1}·Å^{-2}. In the implicit-solvent simulation the C-Co interaction in the implicit-solvent simulation was represented using a hard-sphere potential.

From Figure 4.13, there is some disparity between the explicit-solvent and implicit-solvent peaks which correspond to the separation distances between the cobalt ion and the more distant constrained oxygen atom in the succinate ion. The extent of the differences was examined further by repeating the sampling simulations using lower force constants of 0.05 kcal·mol^{-1}·Å^{-2} at constraint distances of 3 Å. Again there is good agreement between the lower-separation distance peaks in the explicit-solvent and implicit-solvent distributions, as shown in Figure 4.14. However, there are clear differences between the peaks at greater separation distances in the explicit-solvent and implicit-solvent distributions. That such differences arise, even when the configurations sampled are influenced by minimal constraining potentials, indicates

that the interaction potentials utilised in the implicit-solvent sampling simulation cannot precisely replicate the configurations adopted by the cobalt-succinate ion pair when the surrounding solvent is represented explicitly. Nevertheless, the explicit-solvent and implicit-solvent distributions agree sufficiently well as to demonstrate that, when the cobalt ion and the succinate ion are in close proximity, the interaction potentials associated with the implicit-solvent method can reasonably accurately reproduce the configurations sampled by the ion pair when the solvent is represented explicitly.

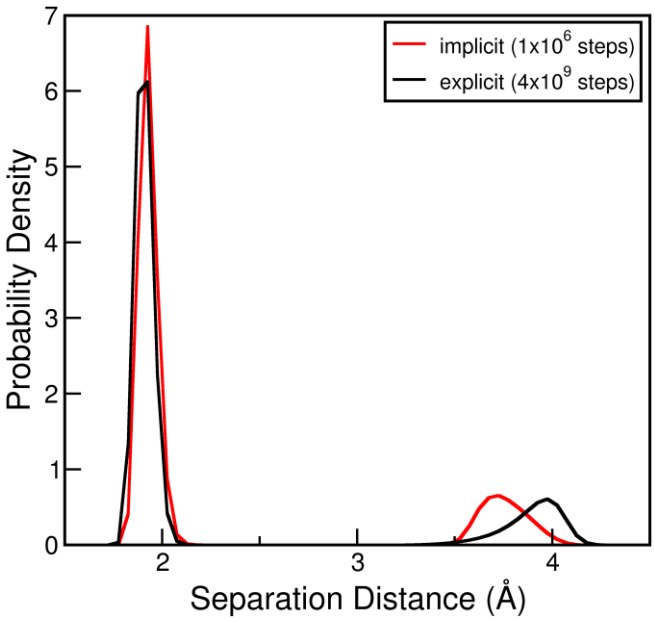

Figure 4.14: Comparison of implicit-solvent and explicit-solvent probability density distributions of the separation distances sampled in simulations involving the succinate-cobalt ion pair constrained, between the cobalt ion and the oxygen atoms at one terminus of the succinate ion, with constraint distances of 3 Å and force constants of 0.05 kcal·mol^{-1}·Å^{-2}. In the implicit-solvent simulation the C-Co interaction was represented using a hard-sphere potential.

The ability of the implicit-solvent potentials to replicate the explicit-solvent sampling was also assessed for the cobalt-succinate ion pair when the cobalt ion and the succinate ion were separated by a greater distance. Because of the favourability of configurations involving the cobalt ion in close contact with one of the oxygen atoms in the succinate ion, to ensure that the ions were further apart, larger constraining potentials were required. Therefore, for implicit-solvent and explicit-solvent representations, in order to obtain distributions of the separation distances between

the ions from simulations in which the cobalt ion was constrained at, and maintained close to, a distance of 4 Å from the oxygen atoms in the succinate ion, a force constant of 8 kcal·mol^{-1}·Å^{-2} was required. The agreement, shown in Figure 4.15, between the resulting explicit-solvent and implicit-solvent distributions indicates that, under the influence of the constraints, the implicit-solvent interaction potentials can reproduce the behaviour of the cobalt-succinate ion pair in an explicitly-represented solvent. However, the value of this comparison is limited by the necessity for a high constraining potential.

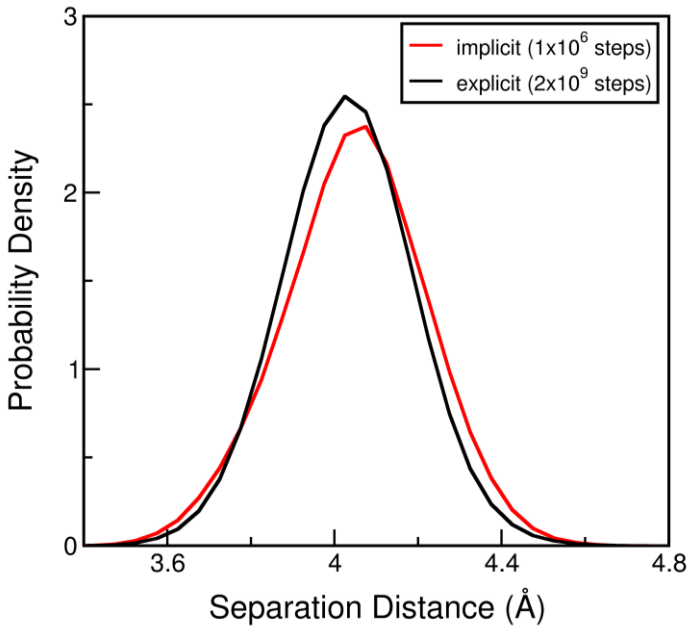

Figure 4.15: Comparison of implicit-solvent and explicit-solvent probability density distributions of the separation distances sampled in simulations involving the succinate-cobalt ion pair constrained, between the cobalt ion and the oxygen atoms at one terminus of the succinate ion, with constraint distances of 4 Å and force constants of 8 kcal·mol^{-1}·Å^{-2}. In the implicit-solvent simulation the C-Co interaction was represented using a hard-sphere potential.

Extending the assessment of the capacity of the implicit-solvent representation to separation distances close to 5 Å required the application of force constants of 5 kcal·mol^{-1}·Å^{-2}. With the configurations sampled by the cobalt-succinate ion pair influenced by such constraints, the resulting explicit-solvent and implicit-solvent distributions differ in terms of the most favourable separation distance between the ions, as shown in Figure 4.16. However, there is good agreement between the ranges

of distances separating the ions in the configurations sampled in the explicit-solvent and implicit-solvent simulations.

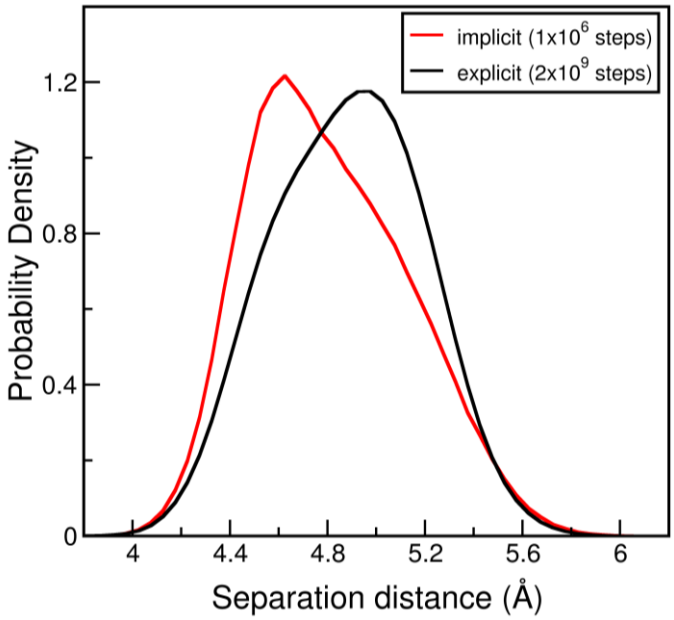

Figure 4.16: Comparison of implicit-solvent and explicit-solvent probability density distributions of the separation distances sampled in simulations involving the succinate-cobalt ion pair constrained, between the cobalt ion and the oxygen atoms at one terminus of the succinate ion, with constraint distances of 5 Å and force constants of 5 kcal·mol^{-1}·Å^{-2}. In the implicit-solvent simulation the C-Co interaction was represented using a hard-sphere potential.

In summary, the implicit-solvent method has been shown to adequately represent the behaviour of the cobalt-succinate ion pair for separation distances close to 3, 4 and 5 Å. For greater separations, however, evaluation of the suitability of the potentials involved with the implicit-solvent representation proved impossible. With the application of constraints which force the cobalt and succinate ions further apart than 5 Å, the overall length of the succinate ion along an axis parallel to the carbon backbone of the ion is such that the succinate ion can rotate to yield configurations in which the cobalt exists in close proximity to an oxygen atom at the opposite terminus of the succinate ion to those involved in the constraints. The ability of the implicit-solvent method to reproduce the explicit-solvent behaviour for equivalent configurations has been assessed previously. The situation could be overcome by employing constraining potentials between the cobalt ion and the oxygen atoms at

both termini of the succinate ion. However, by imposing greater numbers of constraints between the ions, the configurations sampled in the simulations are significantly influenced by the constraining potentials, and the resulting comparison of the explicit-solvent and implicit-solvent methods would have less value as a means of evaluating the capacity of the implicit-solvent representation. Furthermore, for the purpose of simulating the synthesis of MOFs, that the interactions between the self-assembling entities can be accurately represented by utilising the implicit-solvent method is most important at short separation distances. This reflects the fact that, in aligning themselves prior to reaction, the building-block constituents of MOFs are positioned in fairly close proximity. Therefore, the inability to rigorously assess the capabilities of the implicit-solvent representation at large separation distances is not greatly important.

4.2.2.2 Cobalt ion constrained perpendicular to succinate ion carbon chain

Further to the assessment of the capacity of the implicit-solvent method to replicate the explicit-solvent configurational sampling when the cobalt ion was constrained parallel to the carbon chain in the succinate ion, the implicit-solvent method was evaluated for the situation involving the cobalt ion constrained perpendicular to the carbon chain in the succinate ion. A schematic of such an arrangement of the ions is illustrated in Figure 4.7b. Evaluation of the capacity of the implicit-solvent representation to yield the explicit-solvent sampling for configurations involving the cobalt ion approaching the succinate ion perpendicular to the chain of carbon atoms required the application of equal constraints between the cobalt ion and the two carbon atoms in the succinate ion, as shown in Figure 4.17. As before, due to the propensity for the cobalt ion to occupy positions close to an oxygen atom in the succinate ion, when the ions are constrained in this way and a short constraint distance is employed, the ion pair exists in configurations for which the ability of the implicit-solvent method has previously been evaluated. Therefore, with the ion pair constrained between the cobalt ion and the carbon atoms, the initial constraint distance used was 5 Å.

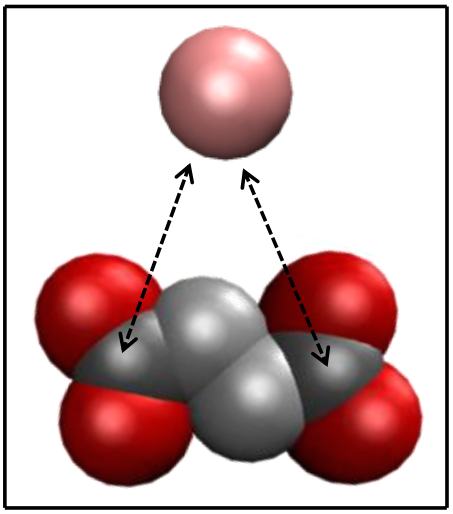

Figure 4.17: The cobalt was constrained perpendicular to the carbon chain in the succinate ion by means of constraints (represented by dashed arrows) between the cobalt ion and the carbon atoms in the succinate ion.

To maintain the cobalt and succinate ions close to a separation distance of 5 Å, a force constant of 10 kcal·mol^{-1}·Å^{-2} was required. The distributions of separation distances, sampled in explicit-solvent and implicit-solvent simulations, between the cobalt ion and all four oxygen atoms in the succinate ion are shown in Figure 4.18. The two peaks evident in the distributions shown in Figure 4.18 correspond to the separation distances between the cobalt ion and, respectively, the closer and further oxygen atoms on the termini of the succinate ion. The correspondence between the peaks in the explicit-solvent and implicit-solvent distributions indicates that, under the influence of the applied constraints, the cobalt and succinate ions occupy similar relative positions in both solvent representations. This good agreement between the distributions resulting from simulations involving implicit and explicit solvent representations indicates that the implicit-solvent effective potentials can accurately replicate the behaviour of the cobalt-succinate ion pair. The value of the comparison is, however, diminished by the necessity for the application of strong constraining potentials. For separation distances greater than 5 Å, to assess the ability of the potentials associated with the implicit-solvent representation to replicate the explicit-solvent behaviour for the cobalt ion approaching the succinate ion perpendicular to the chain of carbon atoms the constraining potential required was of such magnitude compared to the strengths of the potentials at the separation distances involved as to

overwhelm the actual interaction between the ions. Thus, comparisons of the explicit-solvent and implicit-solvent distributions from simulations of the cobalt-succinate ion pair constrained at greater distances would be tantamount to evaluations of the impact of the applied constraining potentials.

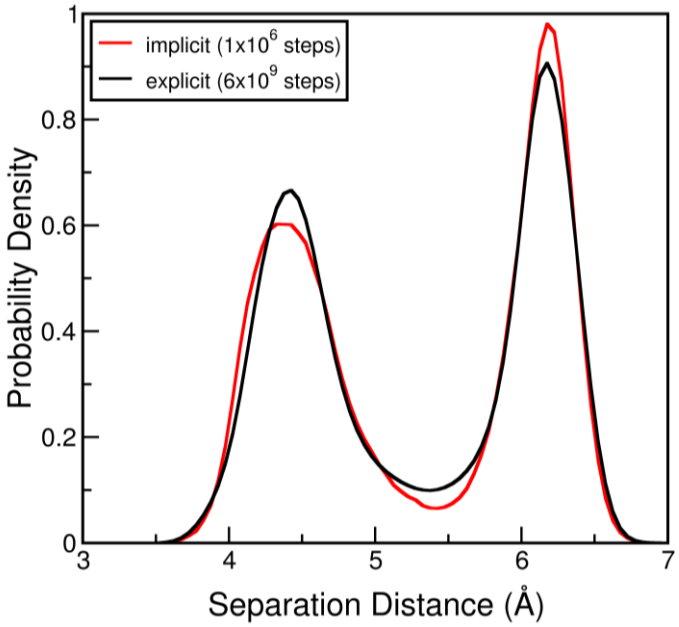

Figure 4.18: Comparison of implicit-solvent and explicit-solvent probability density distributions of the separation distances between the cobalt ion and the oxygen atoms in the succinate ion when the cobalt-succinate ion pair was constrained, between the carbon atoms in the succinate ion and the cobalt ion, at a distance of 5 Å using a force constant of 10 kcal·mol^{-1}·Å^{-2}. The C-Co interaction was represented in the implicit-solvent simulation using a hard-sphere potential.

In summary, in order for the implicit-solvent method to accurately represent the interactions between the cobalt and succinate ions, the carbon-cobalt interaction was represented by a hard-sphere potential. The implementation of this modification to the implicit-solvent method allowed problems associated with the utilisation of an entirely PMF-based approach to be overcome. The evaluation of the capacity of the implicit-solvent representation to replicate the explicit-solvent behaviour of the cobalt and succinate ions was, however, restricted in terms of the range of separation distances that could be examined. Nevertheless, for assessments involving the cobalt ion constrained, respectively, parallel and perpendicular to the chain of carbon atoms in the succinate ion, the implicit-solvent method was shown to have the ability to

accurately represent the explicit-solvent configurational sampling for the extent of the cobalt-succinate separation distances that are most important in simulations of the synthesis of MOFs.

4.2.3 Succinate-succinate configurational sampling

To ensure a detailed examination of the implicit-solvent method with respect to the interactions between succinate ions, the association of the ions was assessed independently for two separate arrangements. With the chains of carbon-containing entities in the succinate ions parallel to one another, the succinate-succinate ion pair was firstly constrained such that one ion approached the other along an axis parallel to the carbon chains, and secondly constrained in order that one ion primarily approached the other along an axis perpendicular to the carbon chains in the ions. The succinate-succinate ion pair arrangements for the two assessments are illustrated in Figure 4.19.

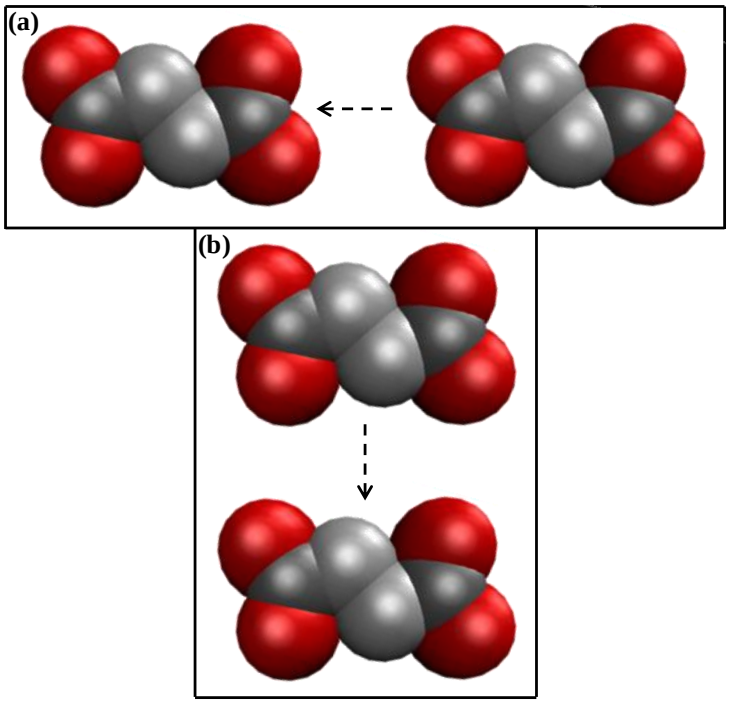

Figure 4.19: The succinate-succinate ion pair was evaluated for, with the carbon chains in the succinate ions in parallel, one succinate ion approaching the other (a) along an axis parallel to the carbon chains in the ions, and (b) along an axis perpendicular to the carbon chains in the ions.

The separate comparisons of the probability density distributions for explicit-solvent and implicit-solvent representations are described in Sections 4.2.3.1 and 4.2.3.2. The ability of the implicit-solvent interactions to represent the behaviour of the pair of succinate ions was determined only for arrangements involving the succinate ions aligned so that the chains of carbon atoms in the ions are parallel, and no assessment was made for arrangements of the succinate-succinate ion pair in which the carbon backbones of the succinate ion were constrained to be perpendicular. This was because succinate ions are unlikely to self-assemble in water to be orientated perpendicular to one another, due to the differing natures of the hydrophobic carbon-containing species and the hydrophilic termini in the succinate ions.

4.2.3.1 Succinate-succinate ion pair approaching along an axis parallel to the carbon chains in the ions

Evaluation of the ability of the implicit-solvent representation to yield the explicit-solvent sampling for configurations involving succinate ions, aligned with parallel carbon chains, approaching along an axis parallel to the carbon backbones required the application of equal constraints, in the form of harmonic potentials, between adjacent oxygen atoms. The constraints applied are shown with respect to the succinate-succinate ion pair in Figure 4.20.

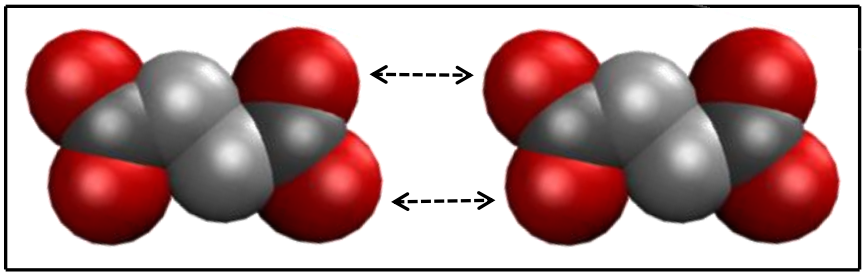

Figure 4.20: The succinate ions were constrained, with parallel carbon backbones, to approach one another along an axis parallel to the carbon chains by means of constraints (represented by dashed arrows) between adjacent terminal oxygen atoms

To compare the explicit-solvent and implicit-solvent methods, the separation distance data from which probability distribution data are constructed were collected between, in one succinate ion, the carbon atom at the terminus involved in the

constraints and, in the other member of the ion pair, the two constrained oxygen atoms. This enabled a thorough analysis of the configurations sampled in the simulations with the respective solvent representations. For the ion pairs investigated so far, the shortest constraint distances for which the implicit-solvent and explicit-solvent representations were compared was 3 Å. However, as a consequence of the repulsion between the oxygen atoms when succinate-succinate ion pair is constrained as shown in Figure 4.20, the force constant required to maintain the separation is of such magnitude as to severely restrict the value of the comparison. Thus, the shortest constraint distance applied with the succinate ions in this arrangement was 4 Å.

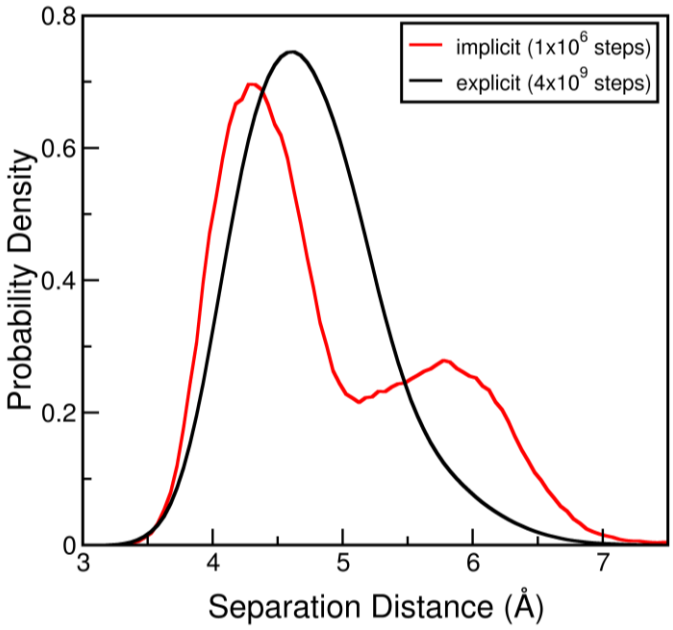

Figure 4.21: Comparison of implicit-solvent and explicit-solvent probability density distributions of the separation distances between a pair of succinate ions constrained, between oxygen atoms, at a distance of 4 Å and using a force constant of 0.5 kcal·mol^{-1}·Å^{-2}.

With neighbouring oxygen atoms in a succinate-succinate ion pair constrained at distances of 4 Å using force constants of 0.5 kcal·mol^{-1}·Å^{-2}, the explicit-solvent and implicit-solvent distributions shown in Figure 4.21 were produced. The application of these constraining potentials between the succinate ions involved in the sampling simulations yields an implicit-solvent distribution in which two distinct peaks are

apparent, whereas the distribution of separation distances produced from the explicit-solvent simulation consists of only one peak. This indicates that with an implicit-solvent representation the favourable relative orientations of two succinate ions differ from the preferred arrangements of the succinate-succinate ion pair in an explicit-solvent representation. Nevertheless, the fact that the maxima in the implicit-solvent and explicit-solvent distributions occur at similar separation distances shows that the most frequently sampled configurations in simulations with implicit-solvent and explicit-solvent representations are not greatly dissimilar. In addition, as shown in Figure 4.21, the implicit-solvent and explicit-solvent distributions have similar baseline ranges. This indicates that, with the different solvent representations, in the configurations possible, under the influence of the applied constraints, the succinate ions are separated by similar distances.

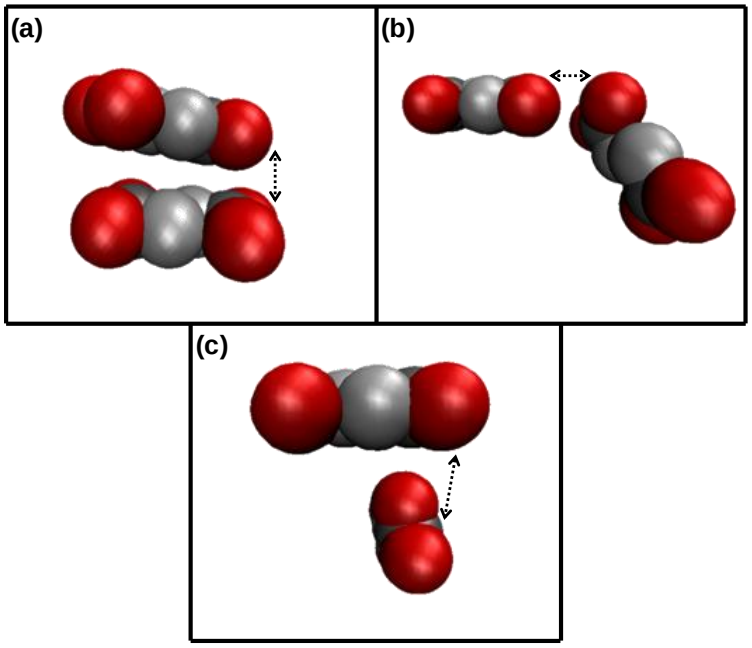

Figure 4.22: Representative succinate-succinate ion pair configurations relating to (a) the peak in the explicit-solvent distribution, (b) the shorter-separation-distance peak in the implicit-solvent distribution, and (c) the longer-separation-distance peak in the implicit-solvent distribution, where arrows reflect the actions of the constraints.

Representations of typical succinate-succinate ion pair configurations that correspond to the peaks in the explicit-solvent and implicit-solvent distributions are shown in Figure 4.22. In this assessment of the capabilities of the implicit-solvent

method, the distributions obtained relate to the association of a pair of succinate ions with constraints applied between the termini. The reasonable agreement between the peak in the explicit-solvent distribution and the shorter-separation-distance peak in the implicit-solvent distribution is reflected in the similarity of both the approach separation distances between the constrained ends of the succinate ions and the relative orientations of the constrained terminal oxygen atoms in the configurations depicted in Figure 4.22a and b. It is apparent, however, that there are significant differences in the overall relative orientations of the succinate-succinate ion pair in these configurations. The clear differences between the approach distances and relative orientations of the constrained ends of the succinate ions in the configurations given in Figure 4.22a and b and that shown in Figure 4.22c, which is representative of those that relate to the higher-separation-distance peak in the implicit-solvent distribution, correspond to the fact that the second peak is not present in the explicit-solvent distribution. Thus, the implicit-solvent method allows the succinate ions to exist, under the influence of the constraints applied, in configurations that are not possible with an explicit representation of the solvent. Nevertheless, whilst the implicit-solvent interaction potentials cannot exactly represent the explicit-solvent behaviour of a pair of succinate ions constrained in this way, the similarities in the distributions resulting from explicit-solvent and implicit-solvent sampling simulations indicate that with an implicit representation of the solvent the explicit-solvent association is, on the whole, adequately well represented.

Because of the repulsion between pairs of the oxygen atoms belonging to different succinate ions, when constraints are imposed between the oxygen atoms, configurations resulting from the application of constraints in the form of harmonic functions centred at 5 Å are more favourable than those sampled at shorter separation distances. Therefore, with a constraint distance of 5 Å, the final probability density distributions can be achieved from sampling simulations involving weak constraints in fewer steps than would be required for simulations involving shorter constraint distances. The explicit-solvent and implicit-solvent distributions resulting from simulations involving the succinate-succinate ion pair constrained, between adjacent oxygen atoms, at separations of 5 Å using force constants of 0.5 $kcal \cdot mol^{-1} \cdot Å^{-2}$ are

shown in Figure 4.23. With the application of these constraints, as with the distributions produced from the succinate-succinate ion pair sampling simulations involving constraints centred at 4 Å, the baseline widths of the implicit-solvent and explicit-solvent distributions shown in Figure 4.23 correspond well. This demonstrates that in the configurations sampled in simulations with, respectively, implicit-solvent and explicit-solvent representations the succinate ions are separated by similar distances. However, whilst the implicit-solvent distribution shows two distinct peaks, the explicit-solvent distribution exhibits a single peak. The fact that the maximum in the explicit-solvent distribution corresponds, in terms of the associated separation distance, to the trough between the two peaks in the implicit-solvent distribution indicates that there are clear differences in the relative arrangements of the succinate ions in the most favourable configurations sampled with explicit-solvent and implicit-solvent representations. Therefore, the implicit-solvent method cannot exactly replicate the explicit-solvent behaviour of the succinate-succinate ion pair constrained as described.

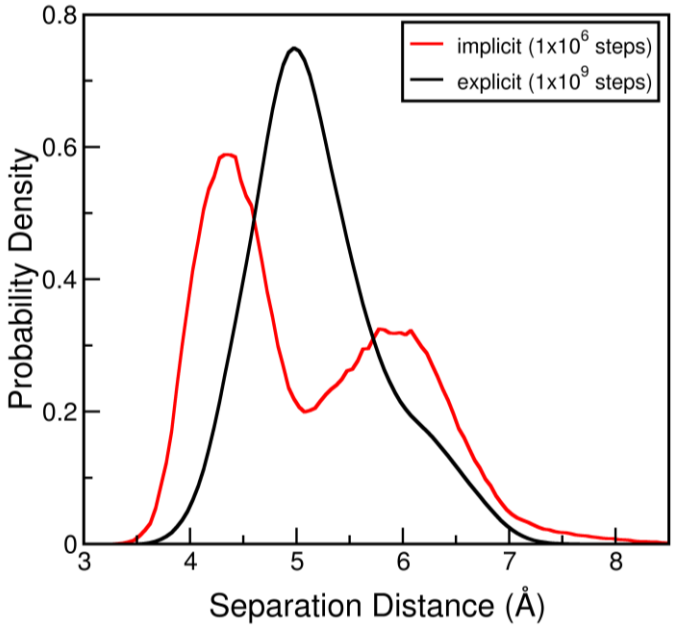

Figure 4.23: Comparison of implicit-solvent and explicit-solvent probability density distributions of the separation distances between a pair of succinate ions constrained, between oxygen atoms, at a distance of 5 Å using a force constant of 0.5 kcal·mol^{-1}·Å^{-2}.

As regards MOF-synthesis simulations of the cobalt succinate system, the succinate ions, which constitute the linker species in the cobalt succinate materials, do not react with other succinate ions. Furthermore, the termini of succinate ions are unlikely to closely approach one another in the absence of a cobalt ion, which is the vertex species in the cobalt succinate materials. Therefore, the fact that the implicit-solvent method is unable to precisely replicate the representative configurational sampling of the succinate-succinate ion pair with an explicit-solvent representation does not detract significantly from the suitability of the implicit-solvent approach as a means of representing the solvent in MOF-synthesis simulations.

For configurations involving the succinate-succinate ion pair constrained, between oxygen atoms, at distances beyond 5 Å, certain pairs of atoms in the ions will be separated by distances greater than the cut-off distances applied in the simulations and, in the implicit-solvent simulation, the atoms are likely to be further apart than the maximum separation distance for which the PMFs were calculated. Since marginal differences in the configurations sampled in the explicit-solvent and implicit-solvent simulations might result in variations in the pairs of atoms between which interactions are considered, the capacity of the implicit-solvent method cannot be assessed reliably. Therefore, with the succinate ions constrained between adjacent oxygen atoms, the explicit-solvent and implicit-solvent representations were not compared for distances beyond 5 Å.

4.2.3.2 Succinate-succinate ion pair approaching along an axis perpendicular to the carbon chains in the ions

The capacity of the implicit-solvent interaction potentials to yield the explicit-solvent sampling for configurations involving a pair of succinate ions, initially aligned with parallel carbon backbones, approaching along an axis perpendicular to the chains of carbon-containing species was assessed. This required the application of equal constraints, in the form of harmonic potentials, between the methylene united atoms in one succinate ion and a carbon atom in the other succinate ion, as shown in Figure 4.24. Constraining the ions in this way limited the configurational sampling possible but still allowed the succinate-succinate ion pair to exist in configurations in which

the ions were not completely parallel. The explicit-solvent and implicit-solvent representations were compared through evaluation of probability density distributions of the separation distances between, in the configurations sampled, the carbon atom involved in the constraints and all four oxygen atoms on the other member of the succinate-succinate ion pair.

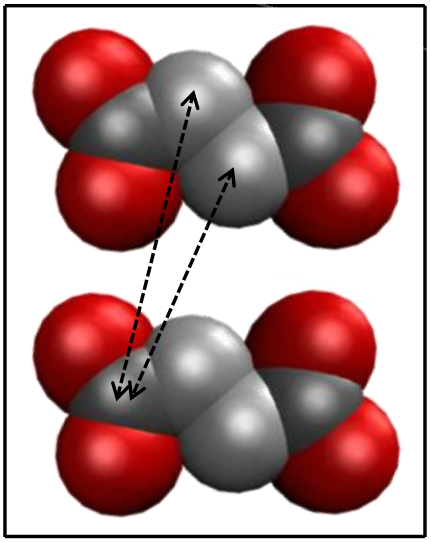

Figure 4.24: The succinate ions were constrained, initially with parallel carbon backbones, to approach one another along an axis perpendicular to the carbon chains by means of constraints between the methylene united atoms in one succinate ion and a carbon atom in the other succinate ion.

With the ions constrained between species located in the backbones of the succinate ions, the shortest constraint distance applied was 4 Å. The distributions of the separation distances between the pair of succinate ions in configurations sampled in explicit-solvent and implicit-solvent simulations in which the ions are constrained via harmonic functions involving force constants of 0.5 kcal·mol^{-1}·Å^{-2} are shown in Figure 4.25. With the application of these constraints there are differences in the ranges of separation distances between the succinate ions in the configurations sampled in explicit-solvent and implicit-solvent simulations. The greater tendency for the succinate ions to exist at further separation distances when the solvent is modelled implicitly indicates that the implicit-solvent method cannot exactly reproduce the explicit-solvent behaviour of the succinate-succinate ion pair.

However, with respect to the most favourable separation distances sampled, there is good agreement between the explicit-solvent and implicit-solvent distributions.

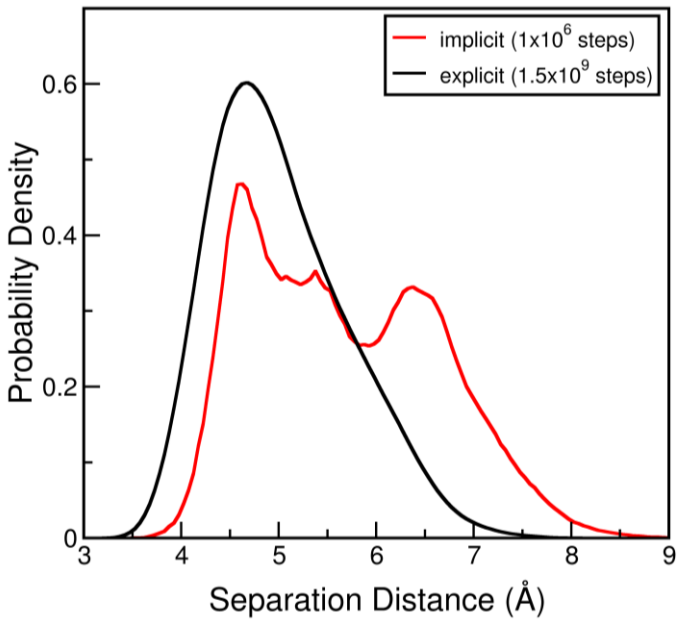

Figure 4.25: Comparison of implicit-solvent and explicit-solvent probability density distributions of the separation distances between a pair of succinate ions constrained, between oxygen atoms, at a distance of 4 Å using a force constant of 0.5 kcal mol^{-1} Å^{-2}.

The implicit-solvent method was further evaluated for the application of 5 Å constraints between species in the carbon backbones of the succinate ions. The explicit-solvent and implicit-solvent probability density distributions produced when a pair of succinate ions are constrained at this separation using force constants of 0.5 kcal·mol^{-1}·Å^{-2} are shown in Figure 4.26. Although the explicit-solvent and implicit-solvent distributions are not entirely dissimilar, the described constraining potentials yield distributions between which a degree of discrepancy is apparent. Additionally, the distribution resulting from the implicit-solvent simulation contains multiple peaks instead of the single distinct peak that is evident in the explicit-solvent distribution.

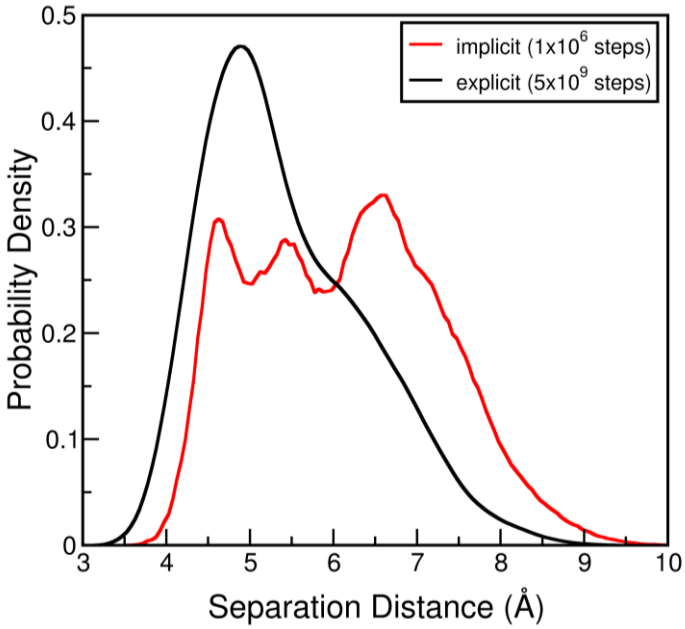

Figure 4.26: Comparison of implicit-solvent and explicit-solvent probability density distributions of the separation distances between a pair of succinate ions constrained, between oxygen atoms, at a distance of 5 Å using a force constant of 0.5 kcal·mol^{-1}·Å^{-2}.

Representative succinate-succinate ion pair configurations sampled in both explicit-solvent and implicit-solvent simulations with the succinate ions constrained at a distance of 5 Å using force constants of 0.5 kcal·mol^{-1}·Å^{-2} are shown in Figure 4.27. The greater baseline range of the implicit-solvent distribution in comparison with the explicit-solvent distribution in Figure 4.26 relates to the fact that, as shown in Figure 4.27b, with the solvent represented implicitly the succinate ions tend to position themselves almost perpendicular to one another. This means that with an implicit-solvent representation the oxygen atoms (in one succinate ion) that are furthest from the other succinate ion are more distant than the equivalent atoms in the explicitly solvated pair of succinate ions. Furthermore, with an implicit-solvent representation there is greater degree of specificity as regards the configurations that can be adopted by the pair of succinate ions. This produces, in the implicit-solvent distribution, several distinct peaks that correspond to the separations associated with specific oxygen atoms, where the peak occurring at greatest separation distances relates to the positions of the two oxygen atoms situated furthest from the other succinate ion in the configurations sampled.

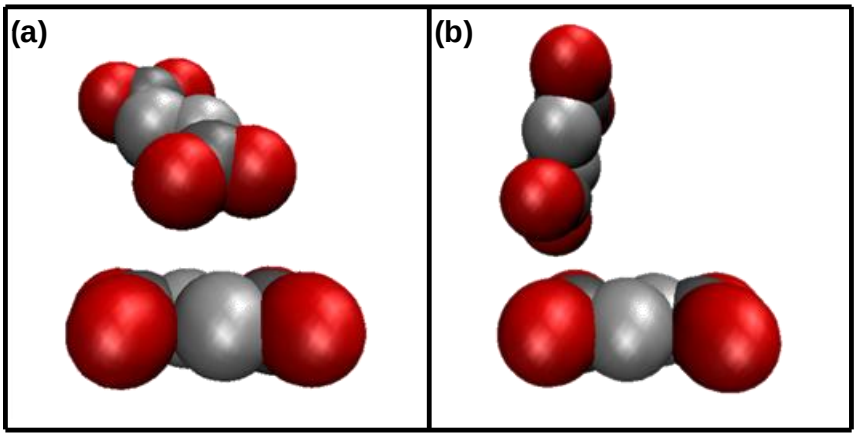

Figure 4.27: Representative succinate-succinate ion pair configurations relating to (a) the explicit-solvent distribution, and (b) the implicit-solvent distribution.

In summary, differences were present to some extent in all evaluations of the ability of the implicit-solvent potentials to replicate the explicit-solvent behaviour for a pair of succinate ions, with the implicit-solvent distributions showing distinct peaks and troughs and much less smooth profiles than the corresponding explicit-solvent distributions. This indicates that the implicit-solvent representation cannot precisely replicate the behaviour of the succinate ions, in terms of the arrangements of the favourable configurations, when the solvent is represented explicitly. This is, perhaps, a result of differences between the environment in which the PMFs are applied (i.e., where the relevant atoms exist within the succinate ions), and the environment in which the PMFs were determined (i.e., where the atoms exist in fully solvated pairs). Nevertheless, the probability density distributions obtained from sampling simulations with implicit-solvent and explicit-solvent representations agree sufficiently well as to demonstrate that the implicit-solvent potentials can adequately approximate the explicit-solvent association of a pair of succinate ions.

4.3 Conclusions

The capacity for the use of PMFs between the constituents of the succinate and cobalt ions to represent the association of the ions has been evaluated through comparison with the explicit-solvent behaviour. For each respective pair of ions, the

explicit-solvent and implicit-solvent behaviour has been assessed by means of sampling simulations, from which probability density distributions of the separation distances between the ions were obtained. The sampling simulations, by necessity, involved the use of constraints between the pair of ions to limit the possible configurations. Thus, the implicit-solvent method has been evaluated as extensively as possible given the restrictions imposed by comparison with the explicit-solvent equivalent.

The overall good agreement obtained in comparisons of the explicit-solvent and implicit-solvent distributions of distances separating a pair of cobalt ions indicates that the Co-Co PMF can accurately represent the explicit-solvent association of the cobalt-cobalt ion pair. The use of only PMFs to represent the interactions between the cobalt and succinate ions was found to prevent the ions from existing in configurations which were favourable with the solvent modelled explicitly. This resulted from the inability of the carbon-cobalt PMF, determined for a pair of isolated species in solvent, to represent certain interactions involving the carbon atom, which is located, in the succinate ion, amongst several other atoms. Replacement of the problematical PMF with a hard-sphere potential enabled the implicit-solvent method to approximate well the explicit-solvent behaviour of the cobalt-succinate ion pair. For the succinate-succinate ion pair, the implicit-solvent interaction potentials between atoms were not capable of precisely replicating the explicit-solvent association but, nevertheless, yielded acceptable approximations of the probability density distributions obtained from sampling simulations involving explicit-solvent representations.

Thus, the interaction potentials associated with the implicit-solvent method have been shown capable of adequately representing the explicit-solvent behaviour of the cobalt and succinate ions for all possible pairs of the ions. Accordingly, the implicit-solvent potentials were considered acceptable as a means of directing the self-assembly of cobalt ions and succinate ions.

CHAPTER 5

Implicit-solvent simulation of the synthesis of MOF materials

Having established in Chapter 4 that the implicit-solvent approach described in Chapter 3 can adequately represent the explicit-solvent behaviour of the primary reactant species in the MOF system under investigation, the next stage in the development of a method for simulating the synthesis of MOFs is to extend the simulation technique employed to represent the organisation of the species in the self-assembly process so as to incorporate a means of allowing the species to react. A simulation method with the potential to accurately predict the formation of MOF structures in relation to imposed reaction conditions must, as a consequence of the reversibility of the self-assembly process by which MOFs synthesise, have the capacity to execute both forward and reverse reactions. These reactions, which involve bond-forming and bond-breaking events, occur between the various different structural entities produced during the progression of the synthesis of a MOF

material. Therefore, a simulation of the synthesis of MOFs must be capable of representing reactions involving the differing species that are generated throughout the course of the simulation and, thus, cannot be identified prior to the commencement of the simulation. In addition, the simulation method must have the ability to yield results within an acceptable time-scale and, ideally, should be easily extended to the study of the synthesis of different MOF systems.

There exist a number of techniques by which reactions might be handled in simulations. These reaction simulation methods can be classified into two distinct categories: quantum mechanical approaches and classical mechanics simulation techniques.[150] Procedures based on the principles of quantum mechanics involve electronic structure theory calculations and, thus, have provision for accurately handling the breaking, formation, or distortion of bonds in simulations of chemical reactions.[151] Approaches involving quantum mechanical calculations include *ab initio* methods such as the Hartree-Fock theory[152, 153] and Møller-Plesset perturbation theory.[154] Such computations involve determination of a molecular wavefunction and can yield highly accurate and consistent results.[155] However, simulations involving *ab initio* calculations are extremely computationally demanding and are therefore limited to small chemical systems.[156] An alternative to standard *ab initio* methods is Density Functional Theory (DFT),[157] which describes the energy of a system as a function of the electron density. Simulations involving DFT calculations can produce reasonably accurate results and are more computationally efficient than conventional *ab initio* schemes.[66] Nevertheless, whilst DFT allows the study of larger systems (containing 100 atoms or more), the computational requirements of DFT techniques are still excessive for many systems.[158]

The high computational demands associated with the application of approaches based solely on quantum mechanics prompted the development of semi-empirical (SE) methods, of which, as reviewed elsewhere,[150, 159-161] there exists a diverse range. SE formalisms involve simplification of the computation necessary in *ab initio* methods and disregard many of the time-consuming terms in the associated calculations. To amend for the approximations introduced, empirical parameters are

incorporated and calibrated against reliable experimental results or highly precise *ab initio* reference data.[162] In addition to being less computationally expensive than techniques based strictly on *ab initio* calculations, SE methods can yield results of acceptable accuracy and have been applied successfully in simulations of a variety of chemical reactions.[163-165] The precision of SE schemes is, however, reliant on both the suitability of the associated approximations and the correctness of the parameterisation,[166] and, whilst SE computations have been found to be highly effective for the description of organic molecules, data calculated for transition metals are generally less accurate.[167]

The second distinct category of methods for modelling reactions, classical mechanical simulations, comprises techniques that have no capacity to directly represent the electron redistribution that occurs in chemical reactions. Nevertheless, such techniques are more computationally efficient than approaches based on quantum mechanics[168] and have demonstrated considerable ability to yield details of the reactivity of chemical systems. Classical simulation methods, such as the reaction ensemble Monte Carlo (RxMC) approach,[169, 170] can permit the elucidation of equilibrium behaviour of chemically reactive systems by considering only the reactants and products of a reaction and not the reaction mechanism. RxMC involves the sampling of predefined forward and reverse reaction events such that the energetics of bond creation and destruction are not required. RxMC has been employed to study a wide range of systems, including those in which the reactions occur in solution and those that involve reactions near solid surfaces.[171] However, because the RxMC method is capable only of predicting the equilibrium properties of reactants defined in advance and has not the wherewithal to create additional species, this approach is not suitable for simulating the synthesis of MOFs. The inability of classical mechanical simulations to describe bond breaking and forming events was addressed by the formulation of reactive force fields, which involve the use of analytical reactive potentials. Such potentials are developed by parameterising experimental reaction energetics data or by executing detailed electronic structure calculations of the pertinent chemical events.[171] A number of distinct reactive force fields have been developed, including ReaxFF[172] and REBO.[173] Analytical reactive

potentials have been employed to study, via molecular dynamics, numerous systems[174] and can enable the simulation of chemical reactions with accuracy almost comparable to that achieved using quantum mechanical methods.[175] However, the development of analytical reactive potentials requires extensive parameterisation and the resulting simulation methods are relatively computationally demanding.[171]

Classical simulation methods can further represent instances of the breakage and formation of chemical bonds by means of the kinetic Monte Carlo (kMC) technique,[176] which can yield insight into the evolution of reactive systems. The kMC simulation method provides a means of reproducing the dynamics of processes occurring over longer time-scales than can be studied using molecular dynamics.[177] By sampling reaction transition processes with probabilities regulated by the ratio of the relevant reaction rate in relation to the rates of all the reaction events possible at that point in the simulation, kMC enables the generation of stochastic trajectories.[178] In general, to obtain the exact evolution, in the real-time domain, of a process, application of this technique requires that all possible reaction events and the associated rates, in terms of real time, are known in advance.[179] These rates depend on the energy barriers between states separated by the reactions and can be obtained from experimental data or derived using computational methods such as transition state theory.[180, 181]

In addition to the two primary categories of reaction simulation methods, quantum mechanical and classical simulation techniques can be combined to produce hybrid approaches, the study of which was instigated by Warshel and Levitt.[182] These techniques involve partitioning the system of interest into regions which are represented using different degrees of approximation. In simulations of reactions, hybrid approaches involve highly accurate quantum mechanical treatment of the atoms directly implicated in the chemical reaction, while the rest of the system, which is assumed not to participate in the reaction and is, therefore, in need of a less detailed description, is described using less precise quantum mechanical methods or empirical molecular potentials.[183] A variety of hybrid methods have been developed, many of which have been applied to study chemical reaction processes.[184] However,

the development of hybrid techniques is generally associated with a significant degree of complexity in terms of the precision of the parameterisation of interactions between the distinct regions.[185]

In striving to develop a computationally practicable simulation method by which the synthesis of MOFs can be represented the method must have the ability to describe the creation and rupture of chemical bonds between priorly unknown species. Furthermore, because a simulation of the synthesis of a MOF material involves a significant number of different entities that can undergo reactions, the MOF-synthesis simulation method must have the capacity to represent systems of quite considerable size and complexity. However, in their standard implementations, none of the outlined existing simulation approaches have the capacity to represent such systems. Therefore, for the purpose of simulating the synthesis of MOFs, formulation of a means of representing the associated chemical reactions was required. The approach utilised in this work is detailed in Section 5.1.

5.1 Simulation Methodology

5.1.1 Simulation Basis

The simulation method utilised for representing the synthesis of MOFs is based on the assumption that the role played by the physical interactions between the building-block constituents of MOFs in aligning the building blocks is of greater importance than the way the reactions between the building blocks are handled. In order to represent the MOF-synthesis process, the simulation method employed in this work involves a form of kMC. As previously described, the standard application of kMC requires the specification, for all possible instances of reaction, of rate constants, which characterise the probabilities with which the different reaction events occur. In this work, the kMC approach utilised reflects the assumption that the rate constants (and, thus, reaction rates) depend primarily on physical factors relating to the alignment of the building blocks, and not on the energetics and dynamics of the reaction. Therefore, within the simulation method developed, the likelihood of a

reaction taking place is dominated by physical interactions, arising from the relative positions and orientations of the individual building blocks. As a result, once the building blocks are properly aligned the probability of reaction is irrelevant. Therefore, the simulation method utilised reflects the fact that, in the self-assembly process by which MOFs are synthesised, the time scale for the physical rearrangement of the building blocks is assumed to be much longer than the time scale for the reaction between the building blocks. However, this assumption that the rate-limiting step is the physical rearrangement of building blocks and not the actual reactions means that the simulation method does not have the capacity to represent the MOF-synthesis process in relation to a quantifiable time.

The formulation of the simulation approach in relation to the outlined assumptions has certain implications in terms of the implementation of the method. In order that the likelihood of a reaction occurring is governed only by the physical interactions between building blocks and not by an associated probability, at any step in the simulation, all species in the simulation have equal likelihood of being involved in an attempt at a reaction move. Furthermore, in assuming that the physical interactions are more important than the means of handling the reactions between the MOF constituents, since all reactions involved in the synthesis of MOFs occur between similarly composed entities, the mechanisms involved in all reactions are assumed to be similar. Therefore, all reactions corresponding to a particular bond-altering process are considered to be accompanied by a certain energy change related to the change in enthalpy associated with the reaction. As regards the suitability of this approach for simulating the synthesis of MOFs, because the primary focus is on the structures of the species that form as a result of a reaction, the actual mechanism leading to the changes in bond connectivity that occur during the reaction event yielding the species is of less concern. Thus, the absence of detailed descriptions of the electron arrangements intrinsic to the reaction mechanism should provide no barrier to accurate simulation of the synthesis of MOFs provided the products of all reactions can be suitably represented.

Crucial to the ability of the simulation method to represent the synthesis of MOFs is the procedure by which the reaction moves are carried out. Details of the algorithm implemented in this work are presented in Section 5.1.2.

5.1.2 Simulation Algorithm

To simulate the synthesis of MOFs, the self-assembly process is represented by executing reaction moves alongside the standard translational and rotational moves employed in MC simulations. Thus, with implementation of the implicit-solvent method developed, these simulations are intended to describe the interaction and association of the self-assembling entities so as to enable the sampling of conformations representative of those generated during the experimental synthesis process and thus determine the formation of the MOF produced under given reaction conditions. A flow diagram outlining the approach developed to simulate the synthesis of MOFs is shown in Figure 5.1.

As shown in Figure 5.1, to represent the synthesis of MOFs, the simulation firstly entails random selection of an entity for which a move is to be attempted. Initially, the simulation involves only entities representative of the organic linkers and the metal vertices but subsequently, as the simulation progresses, the species present include those formed in the reactions that have occurred. Selection at random ensures that, throughout the simulation, all entities in existence at the commencement of any iteration step might be chosen with equal probability. The nomination of a particular entity is followed by random selection of a move type. This appointment of a type of move is performed in accordance with specified ratios describing the relative frequencies with which all of the respective moves available to the chosen entity are to be attempted. By defining equivalent ratios of move types for all of the types of species in the simulation (including those generated in reactions), at any step in the simulation all entities can be considered to have equal probability of being involved in an attempted reaction move. Thus, each step in the simulation involves the trial move, by translation, rotation or reaction, of a randomly selected entity. Trial moves relating to translations and rotations of the selected entity are performed

as described in Chapter 2. The execution of trial moves representing reactions requires further explanation, however.

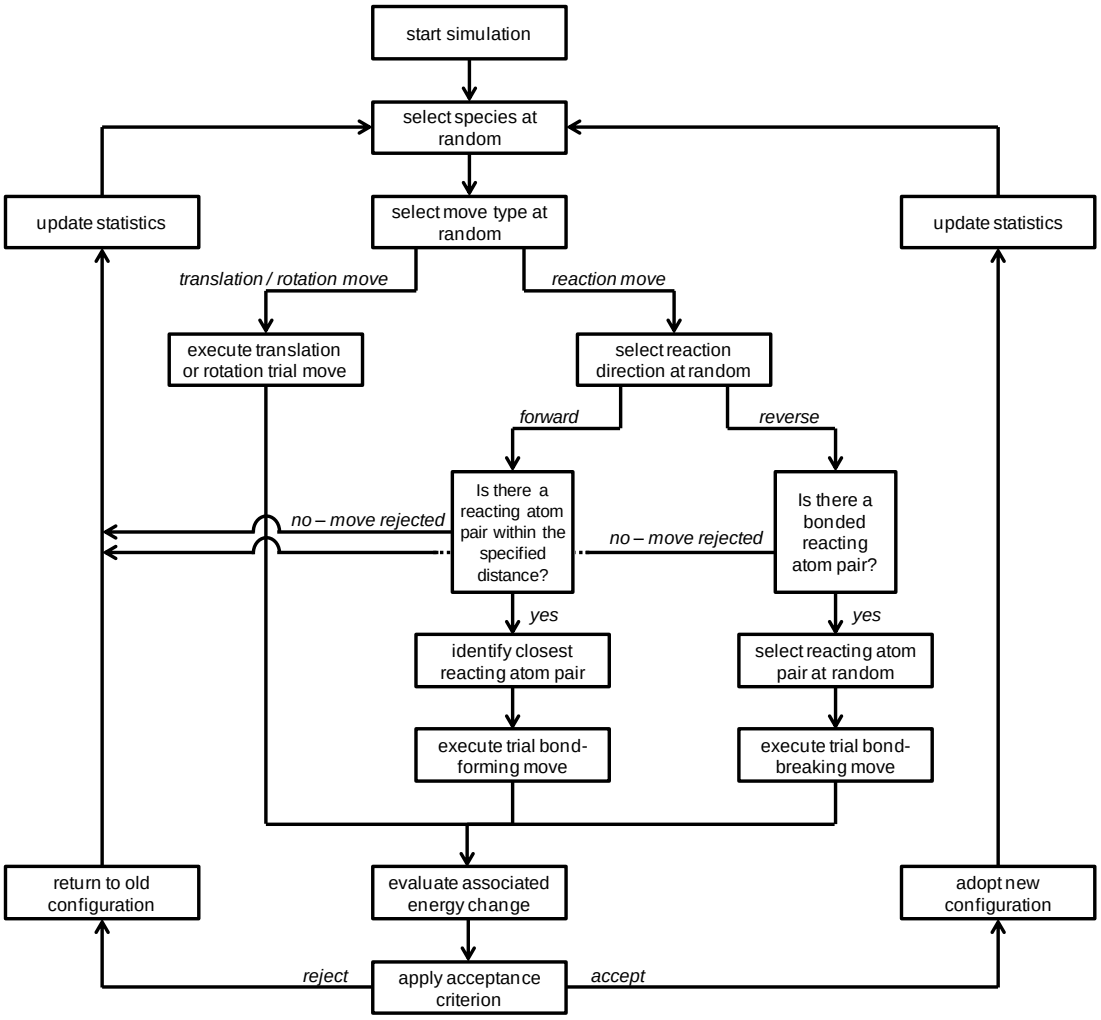

Figure 5.1: Flow diagram showing the steps in the MOF-synthesis simulation algorithm.

The initial operation in undertaking a trial reaction move is to randomly select whether a forward or reverse reaction move is to be attempted. To make certain that the simulation of MOF synthesis pertains as closely as possible to the experimental procedure and does not involve unphysical moves, trial forward reactions proceed only when the constituents of a pair of reacting atoms belonging to different entities exist within a certain separation distance. This reaction separation distance can be specified on a physically justifiable basis through consideration of the form of the reacting atom pair potential of mean force (PMF) employed in the developed

implicit-solvent representation. Thus, both the reacting atom in the selected entity and the corresponding member of the reacting atom pair (which is a component of a second reactant entity implicated in the trial forward reaction) can be identified. However, given the framework-forming natures of the constituents from which MOFs synthesise there is a high likelihood that more than one atom that identifies as a component of the pair of reacting atoms will exist within a particular entity for which a forward reaction is to be attempted. There is a small chance, therefore, that an entity might exist in a configuration whereby more than one of the component atoms are members of pairs of reacting atoms that exist within the reaction separation distance, with the effect that there is not just one single possible forward reaction in which the entity might engage. In this case the forward reaction is attempted between the pair of reacting atoms that are more proximately located. As shown in Figure 5.1, in the circumstance that no pairs of reacting atoms exist within the specified separation distance, the trial forward reaction move is rejected.

As regards trial reverse reaction moves the reaction separation distance is of no relevance because both members of the reacting atom pair initially exist within the same entity. However, in order to carry out a reverse reaction move, the pair of atoms that reversibly react must be identified. For entities containing multiple pairs of bonded reacting atoms, the atom pair involved in the reverse reaction trial is selected at random. When the random selections of the entity and the move type demand that a reverse reaction be attempted with respect to entities (such as those initially present in the simulation) that contain no bonded pair of reacting atoms, the trial reaction move is considered to be rejected.

Having identified the reacting atoms in the entities involved in a trial forward reaction, the move can be attempted. However, there can be no prior knowledge of the structure of the product formed in any reaction. Therefore, the reacting entities are assumed to approximately maintain the relative orientations at which they approach, as directed by the self-assembly process, in the formation of a new product. As a result of the self-assembly process prior to a reaction move, the reactants are considered to be arranged such that their relative positioning is

energetically favourable. Therefore, there is reason to assume that the reactant entities will approximately preserve their spatial structuring when connected to form a single product entity. As previously outlined, to represent the self-assembly process, the simulation method involves, for the purpose of describing intermolecular interactions, the utilisation of pairwise summation of PMFs between the atom constituents of separate entities. As regards intramolecular interactions, as explained in Chapter 2, the model used to describe the succinate ion contains no representation of bond stretching, bond bending, or torsional effects. Furthermore, the specification of parameters to accurately characterise the intramolecular interactions involving bonded atoms for all species generated in reactions would not be feasible. Therefore, all entities in the simulation were represented as rigid structures. Thus, within the simulations, the only intramolecular interactions considered were those between pairs of non-bonded atoms that belong to a single entity but are separated by more than four bonds. For such pairs of atoms existing within the cut-off distance applied, these interactions were described by means of the relevant PMFs.

The attempted forward reaction between two reacting entities involves the formation of a bond between the relevant pair of reacting atoms. Because, for the purpose of developing a method for simulating the synthesis of MOFs, the ability of the simulation technique is being evaluated for existing structures for which the bond lengths between constituent atoms are already known, in this work, the length of the bond created in a forward reaction can be specified according to that occurring between the reacting atom pair in the experimentally derived structure. The formation of the bond in a trial forward reaction serves to link two reacting entities, which are considered to ‘click’ together to form a single product. The opposite occurs in an attempted reverse reaction such that a bond linking two reacted atoms, which exist within a single reactant species, is broken to yield two distinct entities. In addition, there are slight differences in the execution of forward and reverse reaction moves. An attempted forward reaction entails one reacting entity moving closer to the other, whilst approximately preserving an approach direction governed by self-assembly, to form a bond of specified length. Thus, the position to which one reactant must shift to bond with the second reacting entity can be determined.

Furthermore, the relocation associated with a forward reaction is performed in relation to the arrangement of the reacting entities which, as with the surrounding species, exist in positions directed by the prior self-assembly process. Therefore, there is unlikely to be any significant degree of overlap between species as a consequence of this repositioning. However, in the execution of a trial reverse reaction, no precise distance and direction can be defined for a displacement (converse to the coming-together of the reacting entities in a trial forward reaction move) designed to separate the reacting atoms following the bond-breaking event. The only possible route to separating the entities produced in the reaction would be to apply a random displacement as part of a trial reverse reaction move. This would be akin to attempting a translational move in conjunction with the attempt to undertake the bond-breaking incident. Therefore, to avoid the acceptance of a reverse reaction trial being based to some extent on an intrinsic translation, an attempted reverse reaction move consists only of the rupture of the bond between the relevant atoms. As a result, the constituent atoms in the unconnected, individual products generated in a trial reverse reaction move occupy the same positions as the atoms belonging to the associated reactant species.

For both forward and reverse reaction trial moves performed as described, the probability, p, of acceptance or rejection, as with the translational and rotational moves employed in the simulations, is governed by the criterion given in Equation 5.1:

$$p = \min\left[1, \exp\{-\beta(\Delta U)\}\right] \qquad \textbf{5.1}$$

where, as previously stated, $\beta = 1/k_B T$, in which k_B is the Boltzmann constant and T corresponds to the temperature at which the simulation is carried out. The potential change, ΔU, associated with a reaction move represents the difference in energy between the original configuration, which involves the reactant entities, and the configuration that is generated as a result of the trial reaction move and therefore contains a representation of the product species. The energy change associated with

a trial reaction move incorporates an energy term that represents the enthalpy change pertaining to the reaction. As described in Section 5.1.1, the reaction between a particular pair of reacting atoms is associated with a certain enthalpy change, the value of which must be specified prior to the start of the simulation. Acceptance of a reaction move is followed by removal, from the simulation, of the reactants. The simulation thus proceeds from the configuration involving the product species formed in the reaction. Following rejection of a trial reaction move the simulation proceeds from the original configuration containing the reactants.

The outlined procedure constitutes the means by which simulations developed to represent the synthesis of MOFs are executed. As described, the method has the facility to represent reaction moves between the species involved in the formation of MOF structures. The reaction moves are implemented in the simulation method such that there is a requirement for prior specification of the ratio of the move types employed, the reaction separation distance, the enthalpy change associated with the reaction between species, and the length of the bond formed in the reaction between two atoms. With respect to the formation of bonds, in addition to the definition of a bond length between reacting atoms belonging to separate entities, for the simulated entities to yield products with structures that might be realised in the corresponding experimental synthesis, the simulations must involve as realistic as possible a representation of the reactions that occur during the formation of MOFs. This is essential if the simulation method is to have the capacity to predict the MOF structures formed under specified reaction conditions. Therefore, there is a demand for including, in simulations involving accounts of the reactions between entities, a means of ensuring that the species react to form structurally viable products. This requires a more detailed description of the species containing the reacting atoms between which a bond might form during a trial reaction move. In this work, with respect to the creation of bonds between reacting entities, a greater degree of accuracy is realised by utilising the Atoms in Molecules (AIM) method. The underlying theory and resulting capabilities of the AIM analysis are described in Section 5.1.3.

5.1.3 Bond Formation in Reactions

Formulated by R. F. W. Bader, the AIM methodology[186] provides a means of evaluating the electron density distribution associated with a molecule, and thereby yields the capacity to interpret the atoms and bonds, and their relevant attributes, in a molecule. Whilst an overview of the AIM method is presented herein, in-depth descriptions are given elsewhere.[187-189] The AIM analysis involves determining, as a function of the position vector, **r**, the electron density distribution, $\rho(\mathbf{r})$, from either experimental results (obtained using, for example, X-ray crystallography) or theoretical data (derived from, for instance, *ab initio* calculations). The topology of $\rho(\mathbf{r})$ provides a means of partitioning a molecule such that the constituent atoms can be defined via the gradient vector field, $\nabla\rho(\mathbf{r})$, which comprises a collection of gradient paths. As curves that trace the direction of steepest ascent in $\rho(\mathbf{r})$, these paths commence from and terminate at positions in space, known as critical points, where $\nabla\rho(\mathbf{r})$ vanishes. These critical points can be classified as maxima, minima, or saddle points by ascertaining the Laplacian, $\nabla^2\rho(\mathbf{r})$, of the electron charge density distribution. A local maximum is exhibitive of a position where the electron charge density is concentrated with respect to the immediate neighbourhood, and therefore corresponds to a nuclear position. In a gradient vector field in the locality of a nucleus, all of the gradient paths finish at the nucleus, which is, therefore, termed an attractor. The region of space that encompasses all of the gradient paths that terminate at a particular attractor is designated as the basin of the attractor. Thus, the three-dimensional space of a molecular system can be partitioned into atoms, each of which is represented by an attractor and the associated basin.

The properties of the gradient vector field of $\rho(\mathbf{r})$ can further yield information for the description of molecular structure. The critical points in the gradient vector field can be characterised in terms of the eigenvalues (λ_1, λ_2 and λ_3) of the diagonalised $\rho(\mathbf{r})$ Hessian, a 3 × 3 matrix of the second derivatives of $\rho(\mathbf{r})$ with respect to the Cartesian position coordinates of *x*, *y* and *z*. The eigenvalues represent the curvature of the electron density at the critical points, and are negative at a maximum and positive at a minimum. The critical points can be categorised as (ω,σ) in relation to

the rank, ω, which represents the number of non-zero eigenvalues at the critical point, and the signature, σ, which corresponds to the sum of the three eigenvalues. In molecules, four types of critical point are of particular interest: (3,-3), (3,+3), (3,+1) and (3,-1). All curvatures of $\rho(\mathbf{r})$ are negative at a (3,-3) critical point, which is a local maximum and corresponds to a nuclear position. At a (3,+3) critical point all curvatures are positive, and $\rho(\mathbf{r})$ is a local minimum. Because such behaviour is characteristic of the electron density inside a cage, a critical point with (3,+3) classification is called a cage critical point. There are two positive curvatures and one negative curvature at a (3,+1) critical point. This is representative of the topology of the electron density at the centre of a ring of a cyclic molecule and such critical points are, therefore, termed ring critical points. A (3,-1) critical point has two negative curvatures and one positive curvature. This indicates that the point exists between two neighbouring nuclei, signifying the presence of a chemical bond between the nuclei. Such a point is, therefore referred to as a bond critical point. Bond critical points are connected to the adjacent nuclei via atomic interaction lines. Collectively, for a molecular system, the atomic interaction lines define the molecular graph, which generally represents the network of chemical bonds connecting the constituent atoms.

Further, through utilisation of the Laplacian of the charge density, $\nabla^2\rho$, the AIM theory allows detection of reactive sites in a species. Because this quantity allows identification of regions of local electron charge concentration or depletion, the shell structure of an atom can be mapped. So that a charge concentration, or accumulation of electrons, is denoted by a positive value and a charge depletion corresponds to a negative value, for convenience the function $-\nabla^2\rho$ is plotted. Such a plot defines an outermost zone of charge concentration, which is followed by a shell of charge depletion extending to infinity. The outer shell of charge concentration is termed the valence shell charge concentration. In a molecule, within this outer shell there exist regions of bonding charge concentration, which result from bonded interactions, and non-bonding charge concentration, which indicates the presence of lone pairs of electrons, as characterised by the Lewis[190] model. Regions of non-bonding charge

concentration represent sites at which electrophilic attack can occur. Therefore, the AIM analysis has the capacity to indicate, with respect to the reacting atoms in the entities involved in simulations of the synthesis of MOFs, the locations where reaction is possible. Consequently, for the purposes of realistically representing reaction moves, the AIM method can be employed prior to the start of the simulation to identify, for the reacting atoms in the initial species, the regions where bonding is permissible.

5.1.4 Simulation Method – Further Considerations

In the development of a simulation technique there are certain factors, on which the related capabilities are based and which, therefore, merit consideration. Of great importance when implementing a simulation method is that the detailed balance condition is satisfied. As described in Chapter 2, detailed balance implies that when equilibrium is reached the probability of making the transition from one state of the system to another is equal in the forward and reverse directions. There are two key considerations in relation to the detailed balance criterion. The first concern is that all of the transitions occurring in the simulation have a corresponding transition in the opposite direction. Since the magnitudes and directional components of the translational and rotational moves are specified randomly, there is no barrier to the occurrence of equivalent forward and reverse transitions. With respect to the reaction moves implemented in the simulation method, the most exacting route to ensuring that reciprocal forward and reverse reaction moves occur would be to store the details of the reactants from which product species are generated throughout the simulation. This would allow reaction moves to be carried out such that any entity undergoing a reverse reaction could yield, as products, the reactant species involved in the complementary forward reaction. However, the memory requirements of such a simulation would be excessive. In this work, when an entity is involved in a reverse reaction the bond (between reacted atoms) that is to be ruptured is selected at random. Therefore, a species undergoing a reverse reaction move will not necessarily produce entities identical to those which combined in the precise equivalent forward reaction. Nevertheless, since the acceptance of a trial reaction

move is related to the associated energy change, the formation of a very-high-energy species will be similarly unfavourable in both forward and reverse reaction trial moves. Thus, the simulation method is considered to yield similar entities as products of the executed forward and reverse reaction moves. As a result, for the scheme described in this work, for a particular reaction, both reactant and product states are capable of being produced via the transitions (corresponding to reactions) that take place. Therefore, detailed balance is regarded as being upheld if the probabilities associated with the reactions are correctly specified. Similarly, since the trial moves representing translations and rotations involve the species generated in reactions, the fact that both the reactant and product states associated with reaction moves can be produced indicates that there is no barrier to the execution of complementary translational and rotational moves if the related probabilities are accurately defined.

Therefore, the second matter connected with the preservation of detailed balance is that, at equilibrium, equivalent transitions are performed with equal probability. Since all of the entities represented in the simulation are capable of undergoing different reactions, which are assumed to occur with equal probability within a particular step, the initial operation in each step is to select an entity at random. However, this presents a complication with respect to the maintenance of detailed balance because the number of entities involved in the simulation changes, as a result of reaction moves, as the simulation progresses. As previously described, a forward reaction between two reactant entities produces a single product and a reverse reaction yields two product entities from a single reactant. However, because forward and reverse reactions are attempted with equal probability, the number of entities present in the simulation can be considered not to vary significantly provided forward and reverse trial reaction moves are equally likely to be accepted once equilibrium has been reached. This being the case, at equilibrium, in all iterations every one of the entities present can be deemed to be equally likely to be involved in a trial of a particular transition. The preservation of detailed balance requires, however, that, at equilibrium, equivalent forward and reverse transitions between states are performed with equal probability. Since the simulation method has been

demonstrated to attempt corresponding transitions with equivalent probability, the detailed balance condition requires that the acceptance rates of complementary transitions are equal. In the simulation method described in this work the translational and rotational trial moves are implemented such that the magnitudes of the associated relocations are adjusted to ensure a 50 % acceptance rate. The acceptance of reaction moves is dependent on the value of the related enthalpy change. Thus, the extent of the enthalpy change associated with the reaction between a particular pair of reacting atoms should be specified with the intention of achieving equivalent acceptance rates for corresponding forward and reverse reaction trial moves.

By virtue of the fact that a reaction separation distance is defined such that reactions are only permitted to take place between pairs of reacting atoms in separate but closely situated species, the entities involved in reactions must demonstrate some degree of self-assembly. With respect to the self-assembled entities, the acceptance of only trial forward reaction moves (and not attempted reverse reactions) is dependent on the necessity that a relevant constituent pair of reacting atoms is positioned less far apart than the reaction separation distance. However, this requirement, which has the potential to influence the acceptance rate of transitions representing forward reactions, is considered not to have any impact on the preservation of detailed balance. At equilibrium, the self-assembled components in the system are assumed to maintain their level of arrangement to the extent that there is minimal likelihood of an associated trial forward reaction being rejected on the grounds that, in relation to the selected entity, there is no complementary pair of reacting atoms within such distance as to permit an attempt at a forward reaction move. However, in the execution of forward reaction moves, one reacting entity can be displaced by a distance up to that corresponding to the difference between the specified reaction separation distance and the bond length between the reacting atom pair. Therefore, depending on the separation between adjacent self-assembled entities, a forward reaction move has the capacity to disrupt the arrangement so as to render previously feasible reactions impossible. Whilst the implementation of translation moves enables alteration of the displacement distance to ensure 50 %

acceptance, no equivalent practice can be employed in the treatment of forward reaction moves. Therefore, following an accepted forward reaction move there is some likelihood that on a purely distance-related basis, a number of translations are required to recover the extent of self-assembly present (such that equivalent reactions are practicable) before acceptance of the move. Thus, to attempt to ensure that, at equilibrium, trial forward reaction moves of self-assembled species are unlikely to be rejected due to the absence of any appropriate pairs of reacting atoms, consideration can be paid to the magnitudes of the displacements associated with both reaction and translation moves when specifying the ratio defining the relative frequencies with which the different types of moves are attempted.

Therefore, the reaction simulation method formulated can be considered to satisfy the detailed balance condition. This ensures the accuracy of the sampling associated with the simulations and, in addition, because, as reversible reactions, the chemical reactions involved in the synthesis of MOFs are microscopically reversible,[191] the fact that the method enables the execution of simulations representing reactions so as to preserve detailed balance gives further credibility to the results produced. Nevertheless, the test of any simulation method proposed to represent the synthesis of MOFs is the ability of the method to reproduce, under equivalent reaction conditions, the characteristics of the experimentally derived structures. Therefore, the capacity of the approach, of which the underlying assumptions, the operational procedure and requirements, and the underpinning principles have been described, to simulate the synthesis of MOFs was assessed by means of the implementation detailed in Section 5.2.

5.2 Simulation of the Synthesis of MOFs

By implementing the described simulation method alongside the implicit-solvent approach developed, the capacity of the approach to represent the synthesis of MOFs was studied for the cobalt succinate materials for which, as described in Chapter 2, a detailed phase diagram has been defined. As the results presented by way of explaining the development and evaluation of the implicit-solvent method were

obtained at a single temperature (348 K), the assessment of the capabilities of the simulation technique will focus on the cobalt succinate structures that are synthesised at this temperature. Therefore, the capacity of the simulation method was assessed for Phases A and F of the cobalt succinate series. The structures of these phases are introduced in Section 5.2.1.

5.2.1 Cobalt Succinate Structures

As described in Chapter 2, the cobalt succinate materials are formed in the reaction between succinic acid, the organic linker source, and cobalt hydroxide, the metal source, with water as the solvent, such that the reacting species are the succinate ion and the Co^{2+} ion. Phase A,[80] which is the simplest-structured material in the family of cobalt succinates, can be synthesised in isolation, at 348 K, from molar reactant ratios of succinic acid to cobalt hydroxide ranging from 1:1 to 2:1. This phase, shown in Figure 5.2, has a one-dimensional structure and consists of chains of cobalt atoms, which are coordinated by four water molecules, bridged by succinate ions. Therefore, the cobalt atoms in the structure are octahedrally coordinated and form CoO_6 octahedra. The carboxylic groups in the succinate ions are monodentate and, thus, yield free terminal oxygen atoms, which are implicated, along with the water molecules coordinated to the cobalt atoms, in hydrogen bonds that hold the chains together.

At 348 K, Phase F[84] can be synthesised, in isolation, from molar reactant ratios of succinic acid to cobalt hydroxide ranging from 0.2:1 to 0.4:1. The structure of Phase F, which is shown in Figure 5.3, is represented by one-dimensional metal-oxygen-metal chains of CoO_6 octahedra bridged by succinate ligands to form an infinite, two-dimensional sheet. In this phase, the carboxyl groups are bidentate, with the associated oxygen atoms coordinated to different cobalt atoms. Therefore, in the syntheses of both Phases A and F, the reactions between the succinate ion and Co^{2+} ion reactants result in the formation of coordination bonds between the oxygen atoms in the carboxyl groups, and the cobalt atoms. Thus, in simulations of the synthesis of

cobalt succinate materials the reacting atoms are cobalt and the oxygen atoms in the succinate ion.

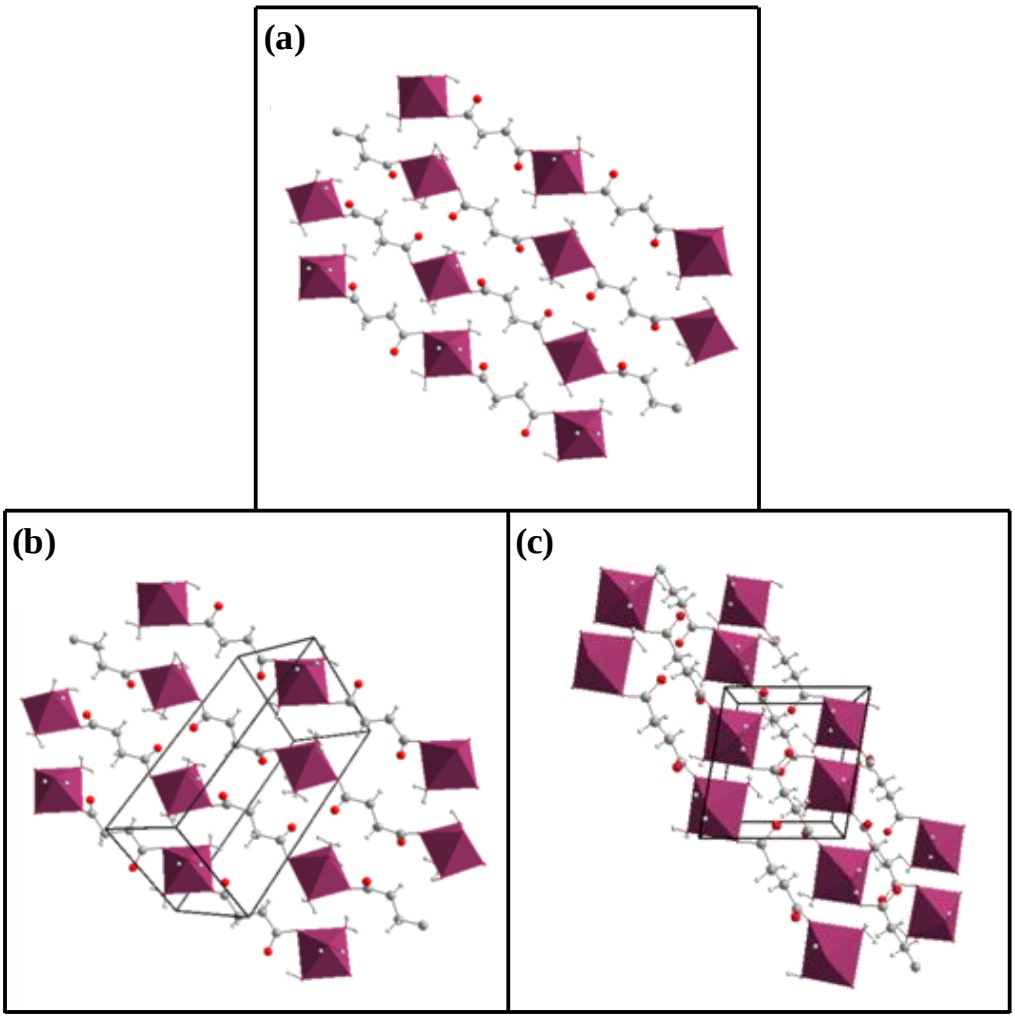

Figure 5.2: The structure of Phase A of the cobalt succinate materials where (a) shows the chains of octahedrally coordinated cobalt atoms bridged by succinate ions, (b) demonstrates the position of the chains within the unit cell, and (c) displays the chains in the unit cell, as viewed in the *ac* plane, in which oxygen atoms are represented by red spheres, carbon atoms are represented by grey spheres, hydrogen atoms are represented by white spheres, the CoO_6 centres are surrounded by purple octahedra, and the edges of the unit cells are delineated in black. The chains shown in these images are held together by hydrogen bonds (which are not illustrated) between the free oxygen atoms in the carboxyl groups and the coordinated water molecules.

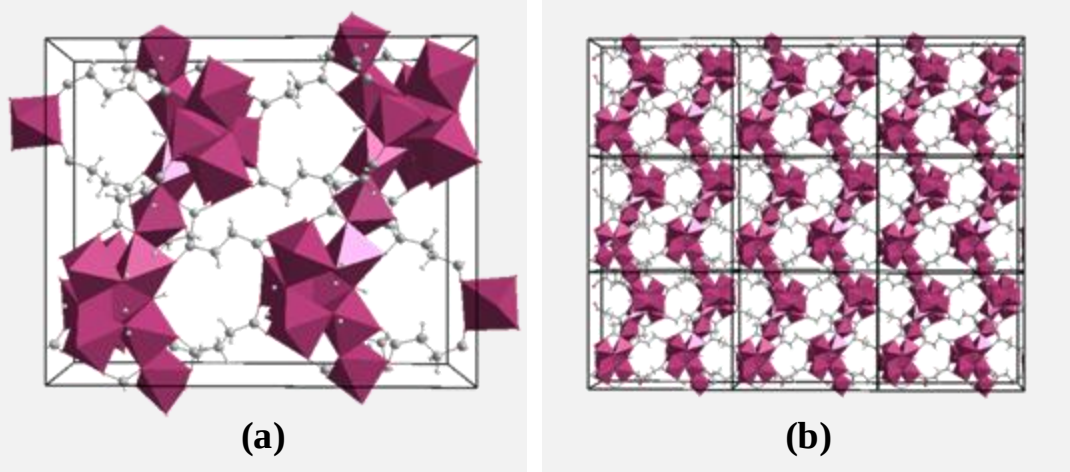

Figure 5.3: Phase F of the cobalt succinate materials as viewed in the *bc* plane where (a) shows the unit cell structure, and (b), in presenting a range of adjacent unit cells, demonstrates the chains of CoO_6 octahedra, which are bridged by succinate ions to form a sheet. In these images oxygen atoms are represented by red spheres, carbon atoms are represented by grey spheres, hydrogen atoms are represented by white spheres, the CoO_6 centres are surrounded by purple octahedral, and the edges of the unit cells are delineated in black.

For the purposes of simulating the experimental syntheses of the materials described, representative ratios of succinic acid to cobalt hydroxide of 1.2:1 and 0.2:1 were utilised, respectively, as the molar reactant ratios under which Phases A and F form. These ratios, along with the specified temperature of 348 K constitute the defined reaction conditions and, thus, correspond to the terms under which the ability of the simulation method to represent the equivalent experimental synthesis is judged. However, evaluating the capacity of the simulation method to represent the synthesis processes of these structures requires, as previously explained, a number of further specifications. The procedures involved in defining the necessary parameters and strictures are outlined in Section 5.2.2.

5.2.2 Simulation Specifications

As previously described, in relation to the implementation of reaction moves, there are a number of requirements that must be specified prior to execution of the simulation method. In the event that a trial forward reaction move can proceed such that a bond forms between the reacting atoms, the length of the bond must be

defined. Details of the bond lengths specified between reacting atoms in Phases A and F of the cobalt succinate materials are outlined in Section 5.2.2.1.

5.2.2.1 Bond Length

Because, in this work, the capacity of the simulation technique to represent MOF synthesis is being evaluated for known phases of the cobalt succinates, the lengths of the bonds formed in forward reactions can be specified in accordance with the lengths of the equivalent bonds in the experimentally synthesised structures. In Phase A the length of the bonds connecting cobalt atoms and the oxygen atoms in the succinate linkers is 2.090 Å.[80] In Phase F of the cobalt succinates the average length of the bonds joining cobalt atoms and the succinate oxygen atoms is 2.092 Å.[84]

These bond lengths can be applied to determine the structure of a new species resulting from the reaction between two separate entities in a trial forward reaction move. As discussed in Chapter 2, to avoid the introduction of inaccuracy, the succinate ion is represented as a rigid structure. For this reason, and because the generation of accurate parameters pertaining to all types of intramolecular interactions in every species formed in all reactions that occur would not be feasible, the entities constructed in reactions are considered to exist as rigid structures. Thus, in the structure of the product of a reaction, the bond connecting the atoms that have reacted is fixed at the length specified.

In carrying out a trial forward reaction move the initial requirement for proceeding is that belonging to the selected species there is an atom within a prescribed reaction separation distance of another atom that represents the corresponding member of the reacting atom pair. The basis on which this reaction separation distance is defined is described in Section 5.2.2.2

5.2.2.2 Reaction Separation Distance

In specifying a separation distance within which the atoms involved in the formation of bonds to yield the structures representing Phases A and F of the cobalt succinates

can be allowed to react, the relevant PMF is examined. The PMF, as determined in Chapter 3, between the reacting atoms (cobalt and oxygen) is shown in Figure 5.4.

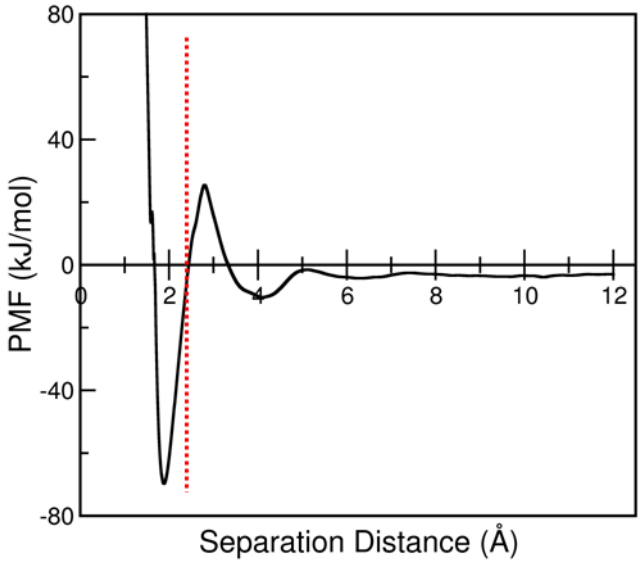

Figure 5.4: PMF between the cobalt ion and the oxygen atom in the succinate ion at 348 K, where the reaction separation distance is indicated by the red dashed line.

As described in Chapter 3, the minimum in the PMF at approximately 1.9 Å represents the contact distance for the reacting atom pair, and the peak at approximately 2.8 Å relates to the transition state between the contact distance and the solvent-shared minimum. The transition state corresponds to the separation distance at which the water molecules are squeezed from between the atom pair. Therefore, within the separation distance relating to the transition state, the approach of one member of the reacting atom pair by the other is favourable and, because the solvent molecules are not positioned between the atoms, minimal reordering of the surrounding solvent molecules is required as the atoms approach.

In this work, in the reaction between a pair of atoms, the bonds that form between reacting atoms are considered to be rigid and maintain the specified length, and there is no associated bond-stretching intramolecular energy term. Consequently, with the formation of a bond between a reacting atom pair, the overall change in energy for the cobalt-oxygen atom pair represents the difference between the energy related to

the separation between the atoms before the reaction (i.e., the PMF value corresponding to the relevant separation distance) and the energy associated with the bonded atoms, which is 0 kJ/mol. Therefore, for the formation of a bond between reacting atoms to be exothermic, the PMF value at the starting reacting atom pair separation must be less than 0 kJ/mol. Thus, for the cobalt-oxygen reacting atom pair, in assuming that a reaction is possible when the solvent does not present a significant barrier to the approach of one reacting atom by the other constituent of the pair, the reaction separation distance was defined as 2.45 Å. In the execution of the simulation method, a trial forward reaction will be allowed to proceed when the selected entity contains an atom that exists within this distance of a corresponding reacting atom belonging to another entity.

In determining whether a trial move involving either a forward or a reverse reaction should be accepted, the energy change related to the reaction move must be ascertained. This energy change incorporates a term representing the enthalpy change associated with the reaction. The means by which the value of the enthalpy change is specified are outlined in Section 5.2.2.3.

5.2.2.3 Reaction Enthalpy Change

Application of the simulation method to the cobalt succinate system requires specification of the enthalpy change relating to the reaction that, in the forward direction, involves the formation of a bond between a cobalt atom and an oxygen atom so as to connect two separate entities. In the absence of data relating to the magnitude of the enthalpy change associated with the reactions involved in the experimental synthesis of cobalt succinates, a suitable value must be defined. As previously described, for the maintenance of detailed balance the enthalpy change should be defined so as to yield equivalent acceptance rates of corresponding attempted forward and reverse reaction moves at equilibrium. Thus, the value representing the enthalpy change is specified with the intention that approximately 50 % of possible trial reaction moves are accepted.

In this work, a particular type of reaction (i.e., the reaction occurring between a given pair of atoms) is considered to be associated with a specific enthalpy change. In the simulation method described, the enthalpy change connected with a reaction constitutes the total energy change within the reacting species due to the formation or rupture of a bond between the reacting atom pair. Therefore, this energy change includes the differences in the interaction energies between all of the pairs of atoms belonging to the entities involved in a reaction. The variations in interaction energies result from changes in the separation distance between pairs of atoms and differences in the interactions that are actually considered. Different atom-atom interactions will be taken into account in the reactant and product species associated with a particular reaction because within a single entity intramolecular interactions are calculated only for pairs of atoms separated by more than four bonds but interactions between all pairs of constituent atoms are represented for the related unconnected entities.

Therefore, to assign a suitable value for the enthalpy change associated with the reactions involved in the synthesis of cobalt succinates, which are generated from initial reactants comprising cobalt ions and succinate ions, a version of the simulation method was applied to represent the association of a pair of the reacting species. Following an equilibration simulation involving only moves representing the displacement and rotation of species, a simulation in which the entities (one cobalt ion and one succinate ion) were allowed to engage in reaction moves in addition to translational and, where applicable, rotational moves (with moves involving translations, rotations and reactions attempted according to the ratio 1:1:0.05). In both simulations the solvent was modelled implicitly using the method detailed in Chapter 3. The reaction moves were implemented so as to proceed only when the distance separating the reacting atoms was less than the specified reaction separation distance of 2.45 Å. In determining the structure produced in the attempted reaction moves, the bond length between the reacting atoms was defined as 2.09 Å. In this version of the simulation method, the reaction enthalpy value was set to 0 kJ/mol. This enabled the collection of data relating to the energy change resulting solely from the reaction between the reactive species. The employment of moves representing displacements and rotations of the species allowed sampling of the energy change

accompanying the reaction for all possible configurations of the pair of initial species. Because the purpose of this simulation was to determine the enthalpy change associated with the reaction, the value was determined only by sampling forward reactions – there can be no means of sampling trial reverse reactions without knowing details of the products within which reverse reactions occur. However, as forward and reverse reactions are assumed to be complementary, the magnitude of the enthalpy change should be equivalent for reactions in both directions.

Probability density distributions of the separation distances between the cobalt ion and the oxygen atoms in the succinate ion are shown in relation to the number of sampling steps in Figure 5.5a. Since no reactions were accepted (because the enthalpy change was specified as 0 kJ/mol in the sampling simulation) the Co-O distances represented in these distributions are characteristic of those occurring when the cobalt and succinate ions approach one another as separate entities. The concurrence of the distributions obtained at different numbers of steps indicates that the simulation adequately samples the configurations of the ion pair. Therefore, the energy change data, from which the enthalpy value is specified, should be representative of that associated with reactions between similar reacting species in all possible configurations. In determining the enthalpy change associated with the reaction, as previously described, the value was specified such that 50 % of the reaction moves sampled in the simulation would have been accepted. Thus, the value of the enthalpy change associated with the formation of a bond between the reacting atom pair was defined as -78.5 kJ/mol.

The cobalt-succinate ion pair sampling simulation was carried out with the intention that the configurations sampled were not governed by the applied ratio of translational, rotational and reaction moves. To have confidence that this was the case, probability density distributions of the Co-O separation distance were determined for equivalent sampling simulations involving different move type ratios. As shown in Figure 5.5b, these distributions, which represent the sampled configurations of the cobalt and succinate ions, correspond for all of the ratios

implemented. Therefore, the enthalpy change specified was not influenced by the imposed move type ratio.

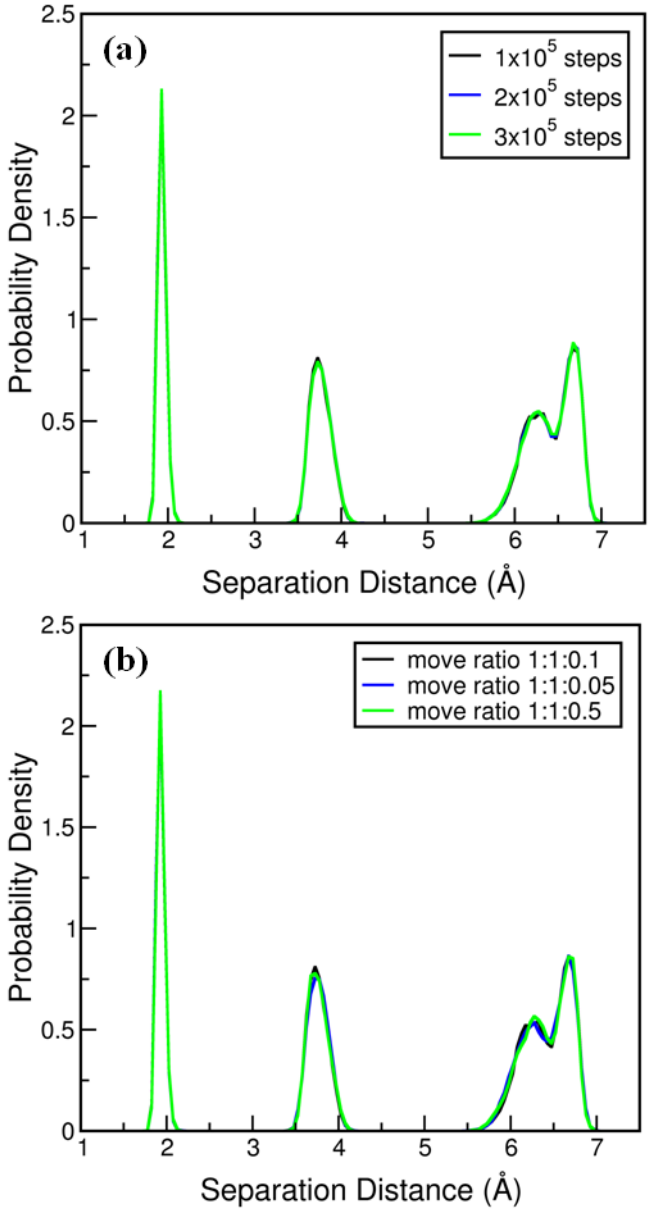

Figure 5.5: Probability density distributions of the separation distance between the cobalt ion and the oxygen atoms belonging to the succinate ion, where (a) shows the distributions sampled after different numbers of steps in the simulation and (b) illustrates the distributions obtained after 1×10^5 steps in sampling simulations with different move type ratios (defined as translation: rotation: reaction).

The execution of simulations representing the formation of the cobalt succinate phases requires specification of a suitable ratio that governs the relative frequencies

with which the different types of moves are attempted. The means by which this ratio is defined for such simulations is described in Section 5.2.2.4.

5.2.2.4 Ratio of Move Types

In specifying a move type ratio the primary objective was that the respective moves should be executed such that the simulation has the capacity to adequately sample different configurations. In addition, as previously described, for the preservation of detailed balance the simulation method requires that the relative frequencies with which different moves are carried out enable the maintenance, at equilibrium, of the level of self-assembly in the system. Thus, specification of the ratio of move types must take into account the capacity for the operation of forward reaction moves to disturb the relative arrangements of the surrounding entities.

The distance shifted by a reacting component involved in a forward reaction relates to the difference between the separation between reacting atoms in the entities as a result of the prior self-assembly and the separation between the reacting atoms due to the formation of a bond (i.e., the bond length). As regards the reactions involved in the production of cobalt succinate materials, knowledge of the typical displacement distance associated with a forward reaction move requires information relating to the characteristic separation, prior to reaction, between the pairs of reacting atom types in the entities involved in the formation of the cobalt succinates. Whilst details of the structures of the entities produced in the synthesis process cannot be known in advance, the interactions between the initial reactants (cobalt and succinate ions) can be assumed to be similar to those between the generated species of which the starting components are constituents. Thus, the typical separation distances between reacting atoms prior to reaction were determined from simulations in which, with an implicit-solvent representation, the cobalt and succinate ions were allowed to interact, and change position by means of moves involving translations and rotations, but not react.

These simulations were carried out to evaluate the characteristic ion-ion separation distances associated with the syntheses of Phases A and F of the cobalt succinates. Thus, the simulations were carried out at the temperature (348 K) and succinate ion to cobalt ion ratios (1.2:1 for Phase A and 0.2:1 for Phase F) corresponding to the reaction conditions under which the respective phases form. The ions were inserted in the simulation cell by means of GCMC simulation, of which a brief outline is given in Chapter 2. In all simulations, cubic simulation cells were employed and, whilst the solvent was represented implicitly, in following the experimental cobalt succinate synthesis procedure that cobalt ions and water molecules should be present in a molar ratio of 1:110, the dimensions of the cells corresponded to those required to yield the experimental water density at the temperature utilised. Distributions of the distances between the cobalt ions and the oxygen atoms in the succinate ions were determined from these simulations once equilibrium had been attained. These distributions, which represent the distances between the pairs of reacting atom types, are shown for two different system sizes for Phases A and F in Figure 5.6. For clarity, only separation distances up to 3.0 Å are presented.

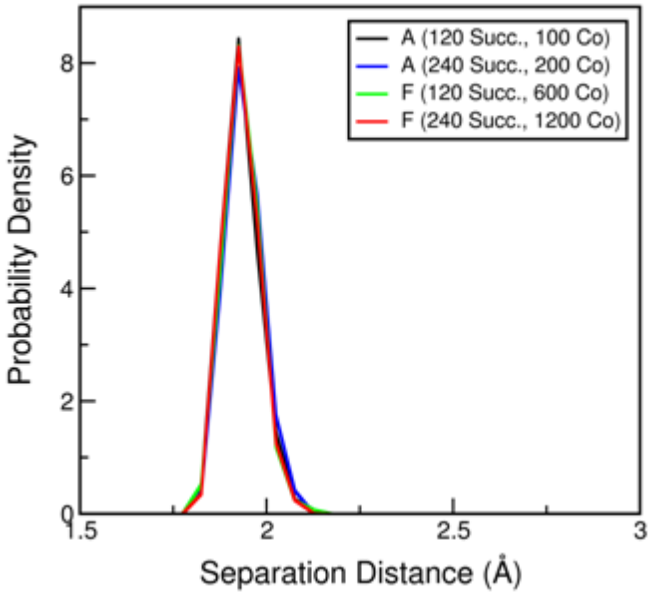

Figure 5.6: Probability density distributions of the separation distance between the cobalt ions and the oxygen atoms belonging to the succinate ions in implicit-solvent simulations in which the ions interact but no reactions are permitted. Distributions are shown for two system sizes, corresponding to different numbers of ions present, for both Phases A and F.

The probability density distributions of the distances between the pairs of reacting atom types indicate that, inside the reaction separation distance, the atoms with the potential to react exist at separations close to the length of the bond (approximately 2.09 Å) that forms between cobalt and oxygen atoms in a forward reaction. Therefore, the displacement associated with an accepted forward reaction move is minimal, and, as a consequence, there will be no requirement for significant repositioning of the surrounding entities to maintain the level of self-assembly present before the reaction. Thus, no consideration need be made of any necessity to specify the move type ratio so as to ensure that, for the purpose of maintaining detailed balance, the entities can regain a prior degree of arrangement. Consequently, as regards the simulation method developed to represent the synthesis of cobalt succinates, the ratio of move types was specified such that translations, rotations and reactions were attempted with equal probability.

For the range of separations distances shown in Figure 5.6, the distribution of distances between the Co-O atom pair indicates that the cobalt and succinate ions exist in particular relative positions. Therefore, in determining the distributions from simulations representing the association of the ions without reaction for the two phases studied, there is merit in examining the arrangements of the cobalt and succinate ions which produce the distributions. The simulations carried out do not permit elucidation of the phase behaviour of the system in which reactions occur because there is representation of only the individual initial reactants and not the species that form when reactions take place. Nevertheless, such simulations can provide some insight into differences in the association of the ions resulting from the application of different synthesis conditions. Employment of the synthesis conditions associated with the formation of Phase A (i.e., a temperature of 348 K and a succinate ion to cobalt ion ratio of 1.2:1) results in the succinate and cobalt ions amassing in a single assemblage. This is evident in Figures 5.7 and 5.8, which, respectively, show configurations arising from simulations of a small system, containing 36 succinate ions and 30 cobalt ions, and a larger system, involving 120 succinate ions and 100 cobalt ions.

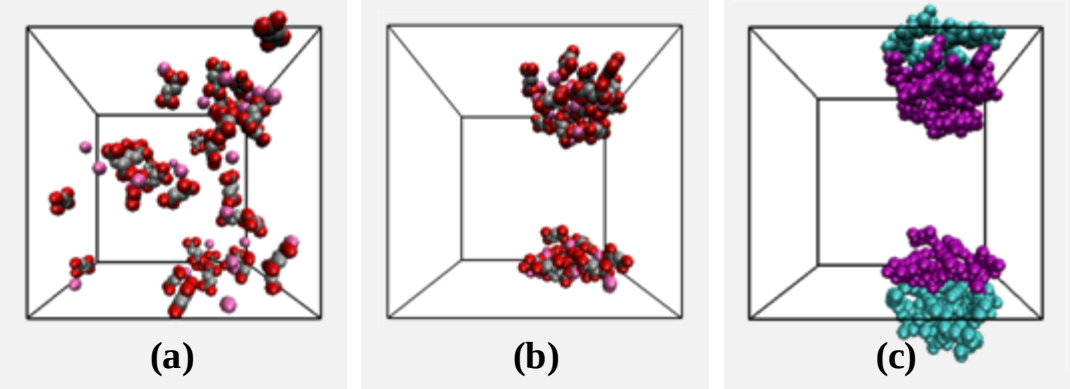

Figure 5.7: Configurations from a simulation representing the association without reaction of 36 succinate and 30 cobalt ions under the synthesis conditions corresponding to the formation of Phase A, where (a) shows the starting positions of the ions and (b) demonstrates an equilibrium arrangement of the ions. In images (a) and (b) cobalt is represented by a pink sphere, oxygen is represented by a red sphere, and carbon and methylene are represented by grey spheres. Image (c) illustrates, by considering periodic replicas of the ions, the grouping of the ions into a spherical cluster. In (c) all constituents of the ions existing in the central unit cell are coloured purple, and all atoms belonging to periodic replica ions are coloured turquoise.

Whilst in the smaller system the cobalt and succinate ions group together into a spherical cluster, as shown in Figure 5.7c, in larger systems, with greater numbers of ions, the configurations sampled in the simulation carried out under the synthesis conditions related to the formation of Phase A reflect the tendency for the cobalt and succinate ions to arrange themselves to form cylinders, as demonstrated in Figure 5.8 for a system containing 120 succinate ions and 100 cobalt ions. Similar behaviour is observed for larger systems but is less comprehensible in presentation. This behaviour is a result of the hydrophobic effect,[192] which leads to the association of organic species in water. As a consequence of the hydrophobicity of their carbon backbones, the succinate ions, of which, as directed by the applied succinate to cobalt ratio, there are a greater number present than cobalt ions, seek to minimise the surface area in contact with the surrounding water, which is represented implicitly in the simulation. The fact that this behaviour is apparent is indicative of the capacity of the implicit-solvent method developed to represent the collective physical characteristics of the ions.

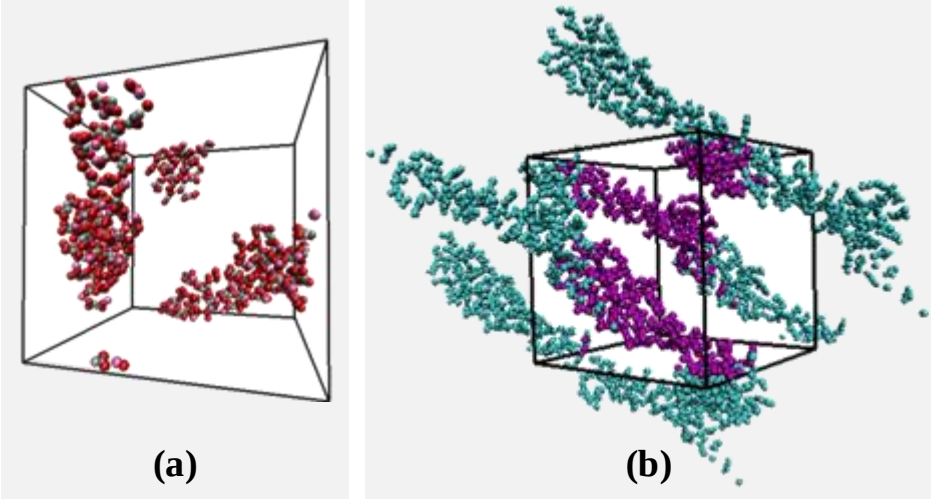

Figure 5.8: Configurations from a simulation representing the association without reaction of 120 succinate and 100 cobalt ions under the synthesis conditions corresponding to the formation of Phase A, where (a) shows an equilibrium arrangement of the ions and (b) demonstrates, by including periodic replicas, the grouping of the ions to form cylinders. In image (a) cobalt is represented by a pink sphere, oxygen is represented by a red sphere, and carbon and methylene are represented by grey spheres. In image (b) all constituents of the ions existing in the central unit cell are coloured purple, and all atoms belonging to periodic replica ions are coloured turquoise.

With the application of the synthesis conditions under which Phase F forms (i.e., at a temperature of 348 K and with succinate and cobalt ions present in the ratio 0.2:1), similar behaviour was apparent for all system sizes. For ease of viewing, the arrangement of the ions is illustrated, in Figure 5.9, for a small system containing 40 succinate ions and 200 cobalt ions. The equilibrium configuration obtained for this system shows that the cobalt ions are dispersed throughout the simulation cell and the succinate ions are positioned in a number of small clusters. This behaviour results both from the tendency of the cobalt ion to be surrounded by six water molecules in an octahedral configuration and from the aforementioned hydrophobic nature of the carbon backbone in the succinate ion. Whilst the arrangements illustrated in Figures 5.7 - 5.9 relate to those sampled in one particular simulation of each respective system, the configurations shown are qualitatively representative of the ion-association behaviour exhibited in each of the multiple (at least three) times that all of the different sizes of systems were simulated under the conditions applied.

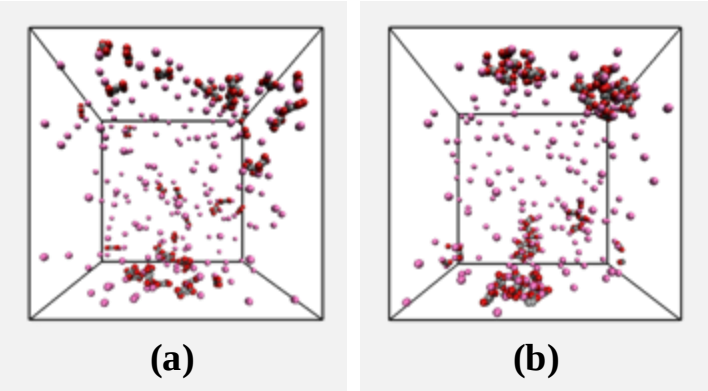

Figure 5.9: Configurations from a simulation representing the association without reaction of 40 succinate and 200 cobalt ions under the synthesis conditions corresponding to the formation of Phase F, where (a) shows the starting positions of the ions and (b) demonstrates an equilibrium arrangement of the ions. In these images cobalt is represented by a pink sphere, oxygen is represented by a red sphere, and carbon and methylene are represented by grey spheres.

In carrying out a simulation of the synthesis of a cobalt succinate phase such that the ions have the capacity to react, in order to adequately represent the experimental reactions information relating to the precise locations of reactive sites is required. As previously explained, this can be achieved by utilising the AIM formulation. The procedures involved in applying the AIM method and the implications as regards simulating the synthesis of cobalt succinate materials are described in Section 5.2.2.5.

5.2.2.5 AIM Analysis

Since all reactions occurring in the formation of cobalt succinate materials take place between cobalt and oxygen atoms, for the purpose of determining, with respect to the reacting entities, the positions of reactive regions, each reaction is considered equivalent to that between the cobalt ion and an oxygen atom in the succinate ion. As the cobalt ion is represented as a spherical entity, there is assumed not to be any circumferential variation in reactivity. Therefore, the succinate ion was analysed using the AIM method to identify any preferentiality with respect to reaction locations. ·

Prior to AIM analysis of the succinate ion, the geometry of the structure was optimised using the (DFT) B3LYP method and the 6-311+G basis set. This approach has been employed successfully to study the conformations of dicarboxylate dianions.[193, 194] As described in Chapter 2, the model employed in this work for representation of the succinate ion involves, through the use of the OPLS-UA force field, implicit inclusion of the hydrogen atoms by means of united-atom methylene groups. However, for the purpose of carrying out the AIM analysis, the optimised succinate ion structure was obtained from an initial all-atom representation in which the hydrogen atoms attached to the carbon backbone were, prior to the optimisation procedure, positioned according to the parameters specified in the OPLS-AA force field.[195] That force fields were combined to produce the starting atomistic representation was of no consequence because the structure was subsequently optimised. The geometry was optimised in the solvent (water) that was employed in the experimental synthesis of the phases belonging to the system of interest. The geometry optimisation and subsequent AIM analysis were performed using the Gaussian09 program package.[196]

The electron density distribution contour plot obtained via the AIM method is shown, alongside the corresponding Laplacian plot, for the succinate ion in Figure 5.10. As shown in Figure 5.10a, the electron density contour lines round the oxygen atoms in the succinate ion are very close together, demonstrating that the electron density changes sharply in this region. In the contour plot of the negative Laplacian of the electron density, the thin crescent-shaped regions round the oxygen atoms indicate that the non-bonding electron density (corresponding to multiple lone pairs of electrons) is smoothed out round the oxygen atoms. The even distribution of the non-bonded charge round the exposed circumference of the oxygen atom (i.e., the part of the oxygen atom where reactions might take place) indicates that all accessible regions on the oxygen atoms are equally reactive and that there is no requirement for specification of orientational preferences of the reacting entities when representing the reaction between oxygen and cobalt atoms in simulations of the synthesis of cobalt succinates.

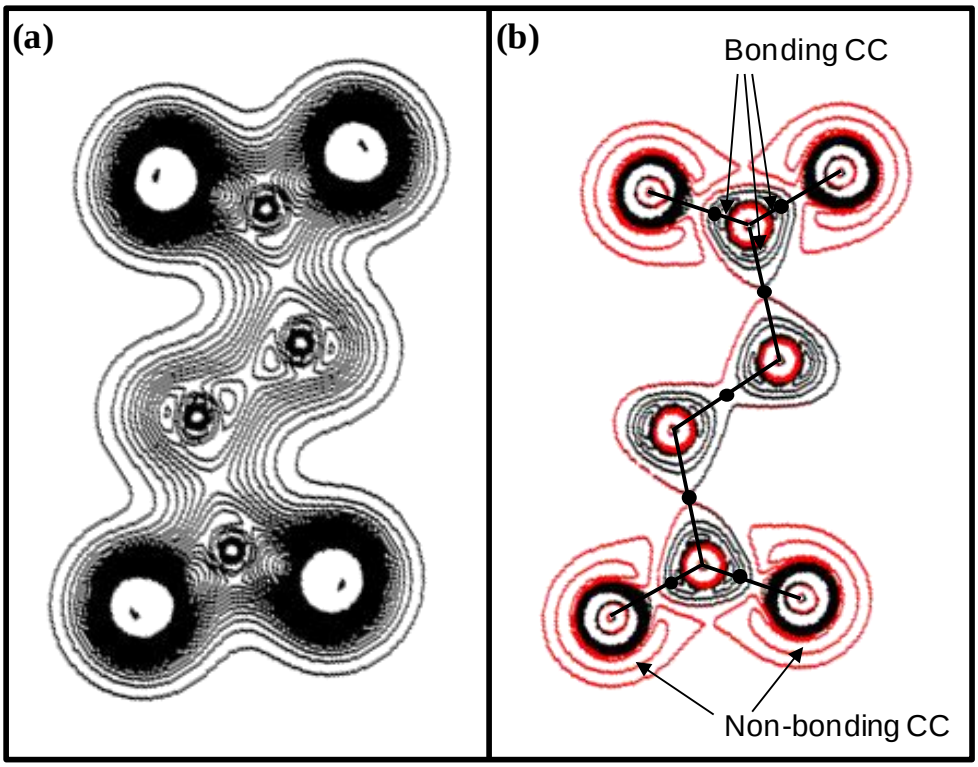

Figure 5.10: Contour plots with respect to the carbon-backbone plane of symmetry in the succinate ion of (a) the electron density distribution in the range 0 to 0.075 e/a_0^3 in steps of 0.025 e/a_0^3 and (b) the negative Laplacian of the electron density in the range -5.0 to 5.0 e/a_0^5 in steps of 0.5 e/a_0^5, where positive contours (which relate to regions of charge concentration) are delineated in red and negative contours (which relate to regions of charge depletion) are delineated in black. Local bonding and non-bonding charge concentrations are indicated by arrows. Atomic interaction lines are depicted as black lines and bond critical points are represented as black circles. The hydrogen atoms bonded to carbon atoms do not appear within the plane of symmetry used.

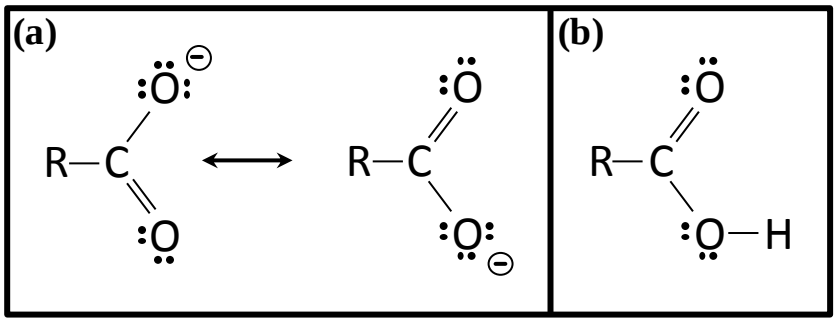

Figure 5.11: The lone pairs of electrons belonging to the terminal oxygen atoms in (a) the resonant structures of a carboxylate ion and (b) a carboxylic acid molecule.

As shown in Figure 5.11a, a carboxylate ion is a hybrid of two resonant structures. Therefore, the electrons can be considered to be delocalised over more than one location. In contrast, in a carboxylic acid molecule, as shown in Figure 5.11b, the

carbonyl oxygen atom has two distinct lone pairs of electrons. Therefore, to have confidence that the AIM method could correctly identify the locations of the non-bonding charge concentration, the same analysis was performed on the succinic acid molecule. For this purpose, an initial succinic acid structure, which underwent geometry optimisation prior to the AIM analysis, was obtained using parameters taken from the OPLS-AA force field.[92] Contour plots of the electron density distribution and the negative Laplacian of the electron density determined for the succinic acid molecule are shown in Figure 5.12.

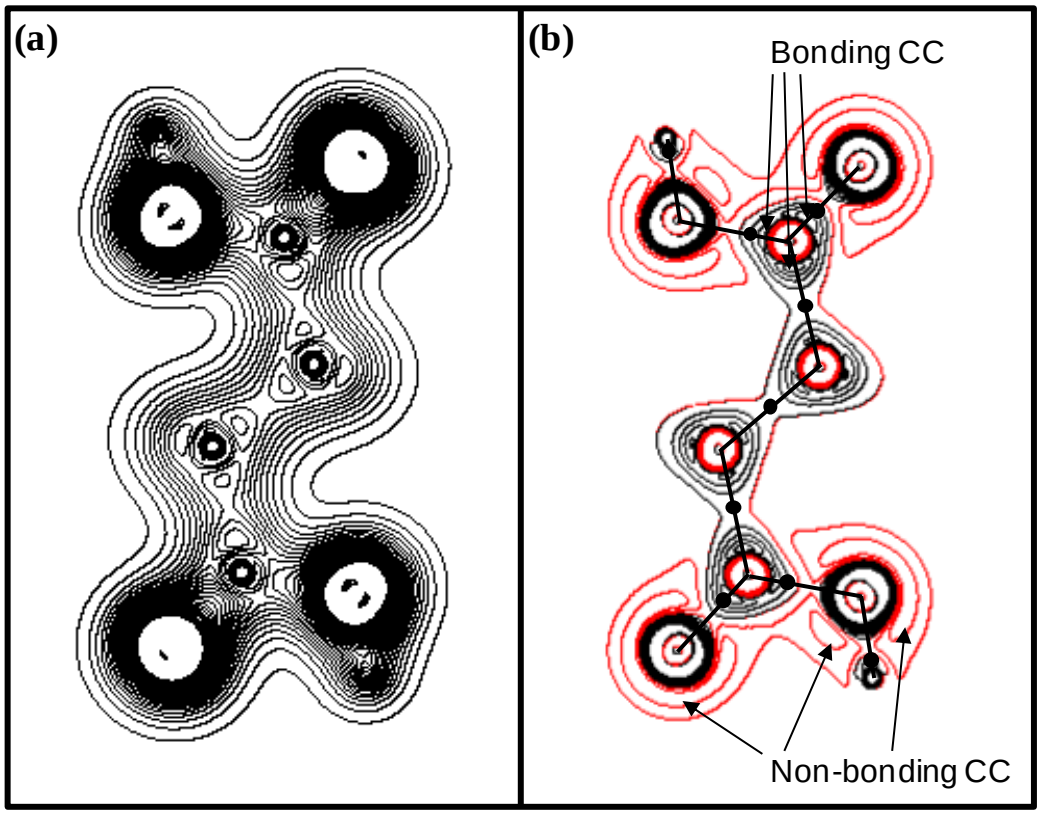

Figure 5.12: Contour plots with respect to the carbon-backbone plane of symmetry in the succinic acid molecule of (a) the electron density distribution in the range 0 to 0.75 e/a_0^3 in steps of 0.025 e/a_0^3 and (b) the negative Laplacian of the electron density distribution in the range -5.0 to 5.0 e/a_0^5 in steps of 0.5 e/a_0^5, where positive contours (which relate to regions of charge concentration) are delineated in red and negative contours (which relate to regions of charge depletion) are delineated in black. Local bonding and non-bonding charge concentrations are indicated by arrows. Atomic interaction lines are depicted as black lines and bond critical points are represented as black circles. The hydrogen atoms bonded to carbon atoms do not appear within the plane of symmetry used.

The contour plot, shown in Figure 5.12b, of the negative Laplacian of the electron density demonstrates that the non-bonding charge concentration is evenly distributed round the carbonyl oxygen atom, which maintains two lone pairs of electrons, and that there are no precisely defined locations at which the associated lone pairs exist. This indicates that with respect to the oxygen atoms in the succinate ion, over which electrons are delocalised, greater resolution of the positioning of the associated lone pairs of electrons would not be expected. Accordingly, the distribution of the electron density determined for the succinate ion was considered to be accurate.

The AIM analysis has, therefore, shown that in simulations of the synthesis of cobalt succinates, the execution of reaction moves requires no consideration of orientational preferences with respect to the reacting entities. Thus, the procedures involved in defining the specifications necessary for execution of the simulation method have been fully described. With the implementation of these parameters, the capacity of the simulation method to represent the synthesis of cobalt succinate materials is examined in Section 5.2.3.

5.2.3 MOF-Synthesis Simulations

To evaluate the capacity of the described simulation method to represent the syntheses of Phases A and F of the cobalt succinate materials, MOF-synthesis simulations were carried out. Execution of the simulations involved the use of a modified version of the Music simulation code.[96] Before commencing the simulations carried out with the aim of representing the syntheses of cobalt succinate materials, starting configurations were generated via GCMC simulation. Thus, for each system studied, succinate and cobalt ions were inserted into the simulation box in numbers consistent with the ratio specified by the relevant synthesis conditions (i.e., ions were present in succinate to cobalt ratios of 1.2:1 and 0.2:1 for Phases A and F respectively). In accordance with the temperature at which Phases A and F are synthesised, all synthesis simulations were carried out at 348 K. All simulations were performed in cubic simulation boxes with dimensions specified such that the inclusion of water molecules in accordance with the cobalt:water ratio of 1:110

would yield the experimental water density at the relevant temperature. In simulations of the synthesis of cobalt succinate materials, the water molecules, which constituted the solvent, were represented implicitly and the interactions between simulated species were evaluated by means of pairwise summation of PMFs (cut-off at a distance of 12.0 Å) between the atom constituents of the species, other than the cobalt-carbon atom pair, for which the interaction was represented by a hard-sphere potential. The relevant PMFs were also employed to represent the intramolecular interactions between pairs of atoms separated by more than four bonds in a particular entity.

As described in Section 5.1.2, the self-assembly process by which the cobalt succinate structures form was represented by allowing the simulated species to undergo moves involving translations, rotations and reactions. In keeping with the specified relation, these moves were attempted in the ratio 1:1:1. Due to the difficulties in assigning accurate intramolecular potentials to species formed as products of reactions, all entities involved in the synthesis simulations were modelled as rigid structures and no moves involving alterations of the internal coordinates of the entities in the simulation were carried out. As outlined in Section 5.2.2.1, reaction moves involved the formation or rupture of a bond between cobalt and oxygen atoms, with the lengths of the bonds between the reacting atoms corresponding to those existing in the experimentally derived structures. Therefore, the lengths of the bonds formed in forward reaction moves were 2.09 Å and 2.092 Å for Phases A and F respectively. Attempts at forward reaction moves were only permitted between different entities containing corresponding reacting atoms separated by a distance less than 2.45 Å. In linking two reacting entities the formation of a bond between reacting atoms was associated with an enthalpy change of -78.5 kJ/mol. When establishing the structure of the product of a forward reaction move, as prescribed by the AIM analysis, there was no requirement for consideration of the orientational preferences of the reactants. Therefore, the structures of the product species formed in reactions were solely directed by the prior self-assembly process that yielded the configuration of the reacting entities.

The procedure and specifications associated with the use of the simulation method to represent the synthesis of Phases A and F of the cobalt succinate materials have been fully detailed in Section 5.2.2. However, the implementation of the simulation method in Music requires a detailed specification of all entities produced in every reaction that takes place. This means that execution of the synthesis simulations is computationally expensive in terms of the accompanying memory requirements. This restricts the extent to which the simulations representing the synthesis of the cobalt succinate phases can proceed. Therefore, in order to assess the capacity of the simulation method to represent the synthesis of cobalt succinate Phases A and F, a balance of the extent of the synthesis process that can be simulated and the possible degree of variation in the products formed was considered. The potential variation in the structures formed in reactions reflects the number of ions present in the simulated system. In systems containing only a small number of self-assembling entities there can be, in comparison with larger systems, relatively few product structures formed in reactions. Therefore, in simulations of small systems, there can be little basis on which to judge the ability of the simulation method to reliably predict the characteristics of the structures produced experimentally. Thus, results are presented for the largest system sizes yielding insightful data. The capacity of the simulation method to represent the synthesis of Phases A and F of the cobalt succinates is described in Sections 5.2.3.1 and 5.2.3.2 respectively.

5.2.3.1 Phase A

The capacity of the simulation method to represent the synthesis of Phase A of the cobalt succinate materials was assessed primarily by studying the system that initially contained 120 succinate ions and 100 cobalt ions in a cubic unit cell with dimensions of 72 Å. For this system, the progression of the synthesis simulation, in terms of the fractions of cobalt atoms coordinated to different numbers of oxygen atoms, is shown in Figure 5.13. The data involved in this representation of the progression of the simulation were obtained by recording, periodically (at a specified interval) throughout the simulation, the numbers of cobalt atoms with each different number of bonded oxygen atoms.

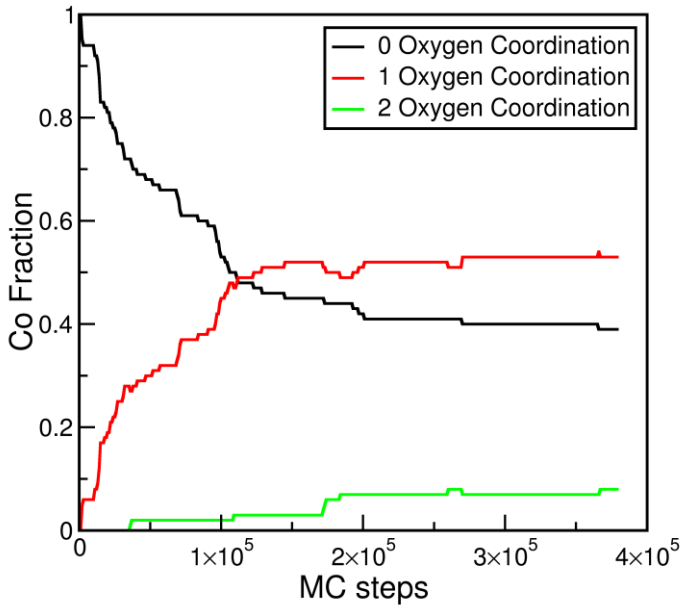

Figure 5.13: Evolution, in terms of MC steps, of the fractions of cobalt atoms coordinated to 0, 1 or 2 oxygen atoms within a synthesis simulation starting from a configuration comprising 120 succinate ions and 100 cobalt ions and involving the reaction conditions corresponding to the formation of Phase A of the cobalt succinate materials.

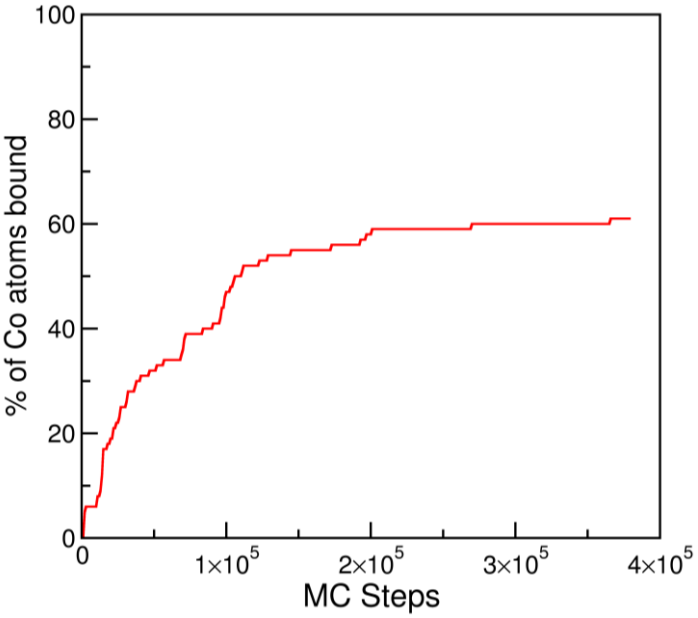

Figure 5.14: Evolution, in terms of MC steps, of the overall % of cobalt atoms that are bound to at least one oxygen atom. This data was obtained from a synthesis simulation starting from a configuration comprising 120 succinate ions and 100 cobalt ions and involving the reaction conditions corresponding to Phase A of the cobalt succinate materials.

Figure 5.13 shows the extent to which the simulation of this system had progressed in a simulation time of approximately two hours. Throughout the extent of the simulation possible for this system, the fractions of cobalt atoms that have coordinated oxygen atoms increase, albeit less rapidly during the final steps than in the initial stages of the simulation. This indicates that within the number of steps simulated, equilibrium has not been attained and the full capacity of the simulation method to predict the structures formed under the relevant reaction conditions cannot be ascertained. In evaluating the cobalt-coordination data, the variation in the overall proportion of cobalt atoms that are bound to oxygen atoms in succinate ions can be studied, as shown in Figure 5.14. As illustrated in this figure, the % of instances in which cobalt atoms are bound to oxygen atoms in succinate ions increases with respect to step number throughout the simulation, including at the final steps represented. This confirms the fact that in the simulation carried out the system does not reach equilibrium, and indicates that, to better evaluate the full capabilities of the approach formulated to represent the synthesis of MOFs, further development of the implementation of the method is necessary. Nevertheless, examination of the simulation progression and the structures of the entities produced in the reactions that have occurred can yield insight into the capabilities of the simulation method. As described previously, Phase A of the cobalt succinates consists of cobalt ions linked by succinate ions to form chains in which the cobalt atoms are coordinated by two oxygen atoms. That, within the extent of the simulated synthesis process, the oxygen-coordination of the cobalt atoms does not exceed two suggests that the simulation method has the ability to represent the experimental structure with sufficient accuracy as to yield the correct coordination preferences of the cobalt ion. There cannot, however, be complete certainty that further coordination would not occur within a longer simulation. Configurations produced in the synthesis simulation of the system studied are shown in Figure 5.15. Comparison of Figure 5.15a, which demonstrates the starting positions of the succinate and cobalt ions, with Figure 5.15b-d, which give representations, from different viewpoints, of the final configuration sampled, demonstrates that the synthesis simulation has resulted in the self-assembly of the entities to form a large cluster. This behaviour corresponds to that observed in the results, shown in Figure 5.8, of the simulation

representing the association without reaction of the equivalent system. As previously described, that the simulated entities arrange themselves in this manner is evidence of the capacity of the simulation to represent the hydrophobic effect.

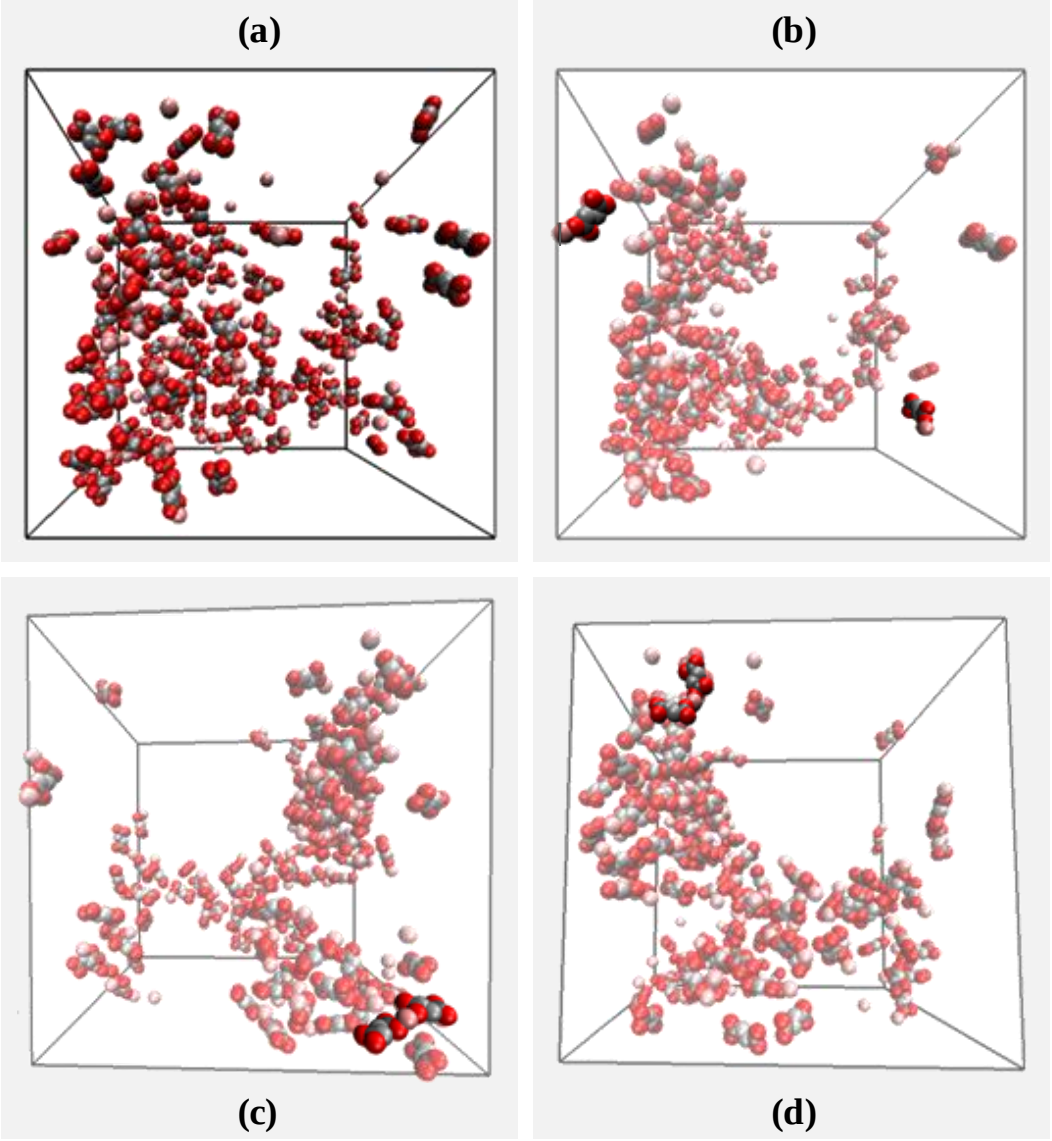

Figure 5.15: Configurations from a synthesis simulation initially comprising 120 succinate ions and 100 cobalt ions and involving the reaction conditions corresponding to the formation of Phase A, where (a) shows the starting positions of the ions. Image (b) highlights the structures of the entities consisting of cobalt atoms coordinated to one oxygen atom in the final configuration as viewed along the z-axis, (c) highlights the structure of a 'correct' short chain in the final configuration as viewed in the negative direction of the y-axis, and (d) highlights the structure of an 'incorrect' short chain in the final configuration as viewed along the y-axis. In these images cobalt is represented by a pink sphere, oxygen is represented by a red sphere, and carbon and methylene are represented by grey spheres.

Typical structures of the primary product entities (as formed in the reaction between a cobalt ion and a succinate ion) are highlighted in Figure 5.15b. The orientation of the ions in these entities is such that there is the capacity for subsequent reactions to result in the formation of chains. This demonstrates that the simulation method has the potential to yield the structures that are expected experimentally. As indicated by Figure 5.13, the final configuration of the simulated system contains a number of entities composed of cobalt atoms coordinated to two oxygen atoms belonging to succinate ions. A structure typical of such an entity is highlighted in Figure 5.15c. That this structure consists of a relatively linear short chain indicates that the simulation method has the ability to accurately predict the characteristics of the underlying structure of the experimentally synthesised materials. For the system initially involving 120 succinate and 100 cobalt ions, examination of all entities composed of a cobalt ion coordinated to two oxygen atoms revealed the existence of the structure highlighted in Figure 5.15d. That this structure is not linear indicates that the simulation method can predict incorrect arrangements. To assess the ability of the MOF-synthesis method to accurately predict structures, characteristics of the entities formed in simulations were compared with those associated with the experimentally synthesised structure of Phase A. Thus, the O-Co-O angles (where the O refers to the oxygen atom that is bonded to the Co atom) in the structures that consist of a short chain composed of bonded succinate-cobalt-succinate were compared with the equivalent angle in the experimental structure. In the MOF-synthesis simulation, the O-Co-O angle in these short-chain structures ranged from 97.5° to 161.4°, with an average value of 126°. The equivalent angle in the structure produced experimentally is 175.7°. Whilst the comparison is based on only the small number of this type of entity produced in the MOF-synthesis simulation, the disparity between the sizes of the angles in the experimental structure and those in the entities formed in the simulation indicates that the MOF-synthesis simulation method does not, at present, have the capacity to yield structures identical to the experimental arrangement of Phase A.

Analysis of a smaller system, which initially contained 36 succinate ions and 30 cobalt ions and was simulated in a cubic unit cell with dimensions of 48 Å,

demonstrates, as highlighted in Figure 5.16, that the simulation method predicts the formation of a non-linear chain of three succinate ions linked by two cobalt ions. In this entity the O-Co-O angles have values of 99° and 130°, and, thus, differ significantly from the equivalent angle (175.7°) in the experimental structure. The greater progress, in terms of the numbers of reactions required to yield the products generated, of the synthesis simulation in which this entity is formed is a result of the smaller system size. However, a system of this scale is less representative of the experimental synthesis.

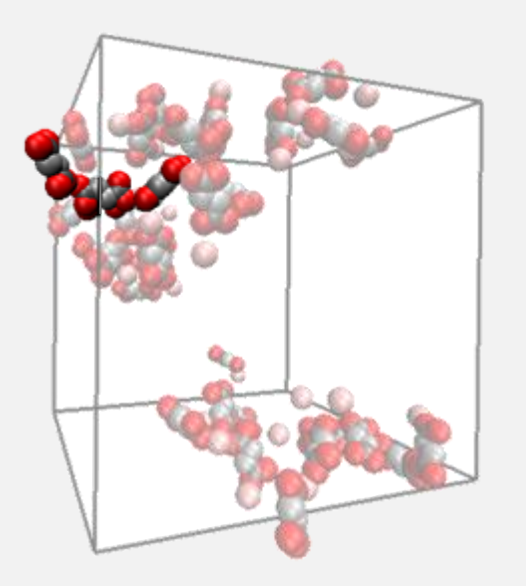

Figure 5.16: The final configuration, in which a chain of three succinate linkers connected by two cobalt ions is highlighted, from a synthesis simulation initially comprising 36 succinate ions and 30 cobalt ions and involving the reaction conditions corresponding to the formation of Phase A. In this image cobalt is represented by a pink sphere, oxygen is represented by a red sphere, and carbon and methylene are represented by grey spheres.

The entities highlighted in Figure 5.15d and Figure 5.16 suggest that, whilst the simulation method has the ability to represent the characteristics of the structures formed experimentally, the method has some difficulty in correctly predicting the precise arrangement of the products of reactions. This is confirmed by the disparity that is evident between the entities produced in the simulation and those in the experimental structure in terms of local structural details such as bond angles. There is a possibility that, with longer simulations, these incorrect structures would

undergo reverse reactions and that the components would then react again to yield acceptably linear chains. However, it is more likely that the incorrect formation of these structures signifies the necessity for the incorporation of some means of allowing the products of reactions to undergo moves that result in alteration of their internal coordinates. This would, however, require the formulation of accurate intramolecular potentials for the entities generated in reactions.

5.2.3.2 Phase F

The capacity of the simulation method to represent the synthesis of Phase F of the cobalt succinate materials was assessed mainly through examination of the system that initially contained 120 succinate ions and 600 cobalt ions. This system was simulated in a cubic unit cell with dimensions of 128 Å. For this system, the progression of the synthesis simulation, in terms of the respective fractions of oxygen-coordinated and isolated cobalt atoms, is shown in Figure 5.17. As before, the data involved in this representation of the progression of the simulation were obtained by recording, periodically throughout the simulation, the numbers of cobalt atoms with each different number of bonded oxygen atoms.

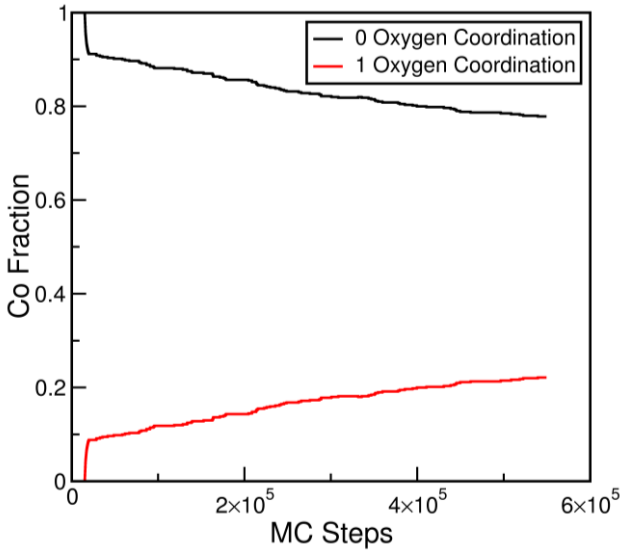

Figure 5.17: Evolution, in terms of MC steps, of the fractions of cobalt atoms coordinated to 0 or 1 oxygen atoms within a synthesis simulation starting from a configuration comprising 120 succinate ions and 600 cobalt ions and involving the reaction conditions corresponding to the formation of Phase F of the cobalt succinate materials.

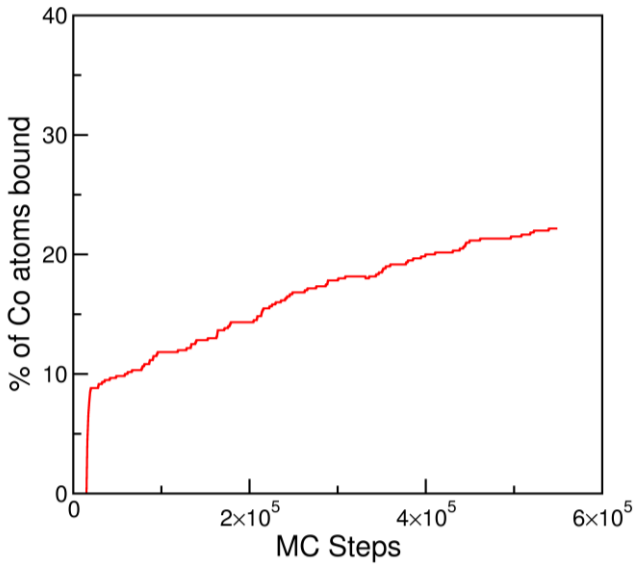

Figure 5.18: Evolution, in terms of MC steps, of the overall % of cobalt atoms that are bound to at least one oxygen atom. This data was obtained from a synthesis simulation starting from a configuration comprising 120 succinate ions and 600 cobalt ions and involving the reaction conditions corresponding to Phase F of the cobalt succinate materials.

Figure 5.17 shows the extent to which the simulation of this system had progressed in a simulation time of approximately four hours. The sharp increase in the fraction of cobalt atoms coordinated to oxygen atoms shortly after the start of this simulation indicates that a certain degree of arrangement of the initial ions is required before any reactions can be accepted. Beyond this point, the fraction of coordinated cobalt atoms increases gradually until the final step of the simulation. This demonstrates that, as with the simulation carried out under the influence of the synthesis conditions associated with the formation of Phase A, within the number of steps carried out in the simulation, equilibrium was not reached. In the simulation carried out under the influence of the reaction conditions associated with the formation of Phase F, there are no cobalt atoms bonded to more than one oxygen atoms. Therefore, in evaluating the cobalt-coordination data, the variation in the overall proportion of cobalt atoms that are bound to oxygen atoms in succinate ions, as shown in Figure 5.18, is equivalent to the evolution, with MC steps, of the fraction of cobalt atoms coordinated to one oxygen atom. Thus, evaluation of the % of instances in which cobalt atoms are bound to oxygen atoms only upholds the fact that the simulation of the system does not evolve to equilibrium, and that, to more thoroughly assess the

capacity of the MOF-synthesis simulation method, further development is required. As before, the capabilities of the simulation method were assessed by examining both the simulation progression and the structures of the entities generated in reaction moves. As previously described, the structure of Phase F of the cobalt succinate materials is composed of chains of CoO_6 octahedra linked by succinate ions such that the carboxyl groups are bidentate. Therefore, in this phase, all oxygen atoms in the succinate ions are coordinated to different cobalt atoms. In being bridged by succinate linkers, the cobalt atoms are generally coordinated to two oxygen atoms belonging to different succinate ions. In Phase F, for the most part, where two succinate ions are coordinated to a particular cobalt atom the succinate ions are positioned such that their carbon backbones are approaching being parallel in alignment.

Upon examination of the structures of the entities produced in the synthesis simulation starting from a configuration containing 120 succinate ions and 600 cobalt ions, the succinate ions were found to be highly coordinated by cobalt ions such that, typically, two or three of the four oxygen atoms in each succinate ion were bonded to cobalt atoms. Typical structures, from the final configuration of this simulation, of entities comprising numerously coordinated succinate ions are shown in Figure 5.19. That such entities are generated in the reactions that take place during the simulation carried out under the reaction conditions associated with the formation of Phase F is evidence of the ability of the simulation method to represent the synthesis of Phase F of the cobalt succinates to the extent of being able to correctly predict characteristics of the experimentally produced structure. Furthermore, that structures consisting of succinate ions with three coordinated cobalt ions are produced in the simulation carried out under the Phase F synthesis conditions but not in the simulation carried out with respect to the reaction conditions associated with the formation of Phase A suggests that, under different reaction conditions, the simulation method has the ability to predict different structures.

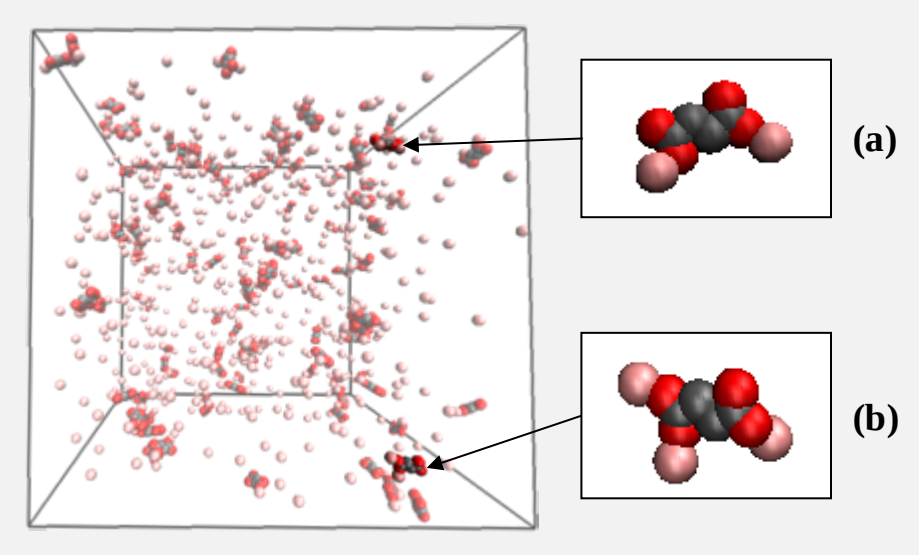

Figure 5.19: The final configuration from a synthesis simulation initially comprising 120 succinate ions and 600 cobalt ions and involving the reaction conditions corresponding to the formation of Phase F. Given in (a) a succinate ion with two coordinated cobalt atoms and (b) a succinate ion with three coordinated cobalt atoms are enlarged versions of the structures highlighted in the final simulation configuration. In these images cobalt is represented by a pink sphere, oxygen is represented by a red sphere, and carbon and methylene are represented by grey spheres.

As shown in Figure 5.17, within the extent of the synthesis simulation that initially contained 120 succinate ions and 600 cobalt ions, no cobalt ions were coordinated to more than one oxygen atom. However, such entities, composed of two succinate ions coordinated to a single cobalt atom, were generated in the equivalent simulation in which initially 100 succinate ions and 500 cobalt ions were present. This system involved the use of a cubic unit cell with dimensions of 120 Å. Within a structure typical of the entities composed of doubly coordinated cobalt ions, as shown in Figure 5.20, the relative positioning of the two succinate ions is similar to that expected in such arrangements evident in the experimentally synthesised structure of Phase F. Indeed, in terms of local structural specifics, the O-Co-O angle (where O refers to oxygen atoms bonded to the Co atom) of interest in the experimental structure is 90.2°, and the average of the equivalent angles in the entities formed in the MOF-synthesis simulation is 113.6°. Whilst this comparison is based only on the small number of examples, in the MOF-synthesis simulation, of the formation of this succinate-cobalt-succinate entity, the reasonable agreement between the values from simulation and experiment demonstrates the capacity of the simulation method to

predict features and local structural details of the structure of Phase F. However, as with the entities generated in simulations carried out under the reaction conditions relevant to Phase A, to achieve greater accuracy in predicting the precise arrangements of the entities formed in reactions would require the implementation of moves enabling modification of the internal coordinates of the entities. Nevertheless, that the structure of the entity displayed in Figure 5.20 differs from the typically chain-like structures produced in simulations involving the reaction conditions associated with the synthesis of Phase A shows that the simulation method described in this work has the ability to predict different structures under the influence of different reaction conditions.

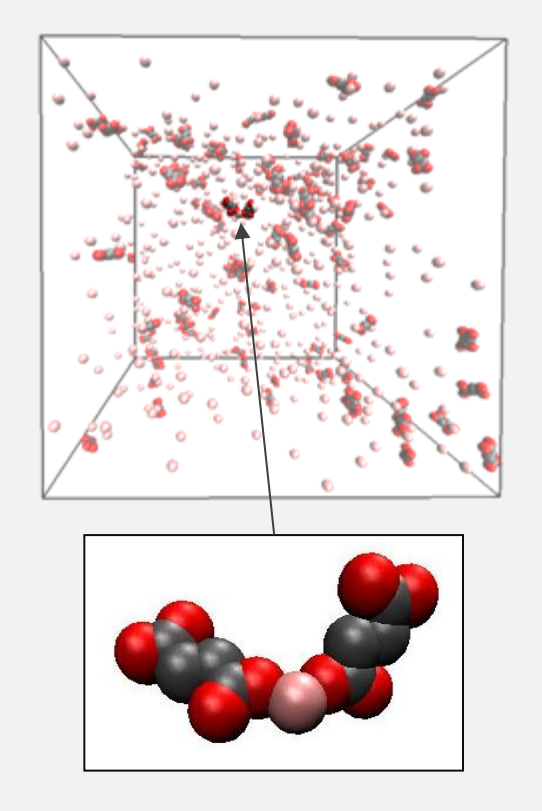

Figure 5.20: The final configuration from a synthesis simulation initially comprising 100 succinate ions and 500 cobalt ions and involving the reaction conditions corresponding to the formation of Phase F. An entity composed of a cobalt atom coordinated to two oxygen atoms within succinate ions is highlighted (and shown in greater detail below the unit cell). In this representation cobalt is represented by a pink sphere, oxygen is represented by a red sphere, and carbon and methylene are represented by grey spheres.

5.3 Conclusions

With the aim of developing a means of representing the synthesis of MOFs, a kinetic Monte Carlo simulation method has been formulated. The capacity of the simulation method to predict the structures formed under different reaction conditions has been assessed through application to the study of the formation of two cobalt succinate phases. Whilst the implementation of the method is limited by the associated computational requirements, the simulation results demonstrate the ability of the method to represent the syntheses of Phases A and F of the cobalt succinate materials. By incorporating, from the equivalent experimentally known structures, only the length of the bond formed between reacting atoms, the simulation method has been shown to have the capacity to correctly predict characteristics of the experimentally synthesised structures of the studied cobalt succinate phases. In so doing, that the simulation method has the ability to yield differing structures under the application of different reaction conditions has been demonstrated.

Whilst the simulation method developed has been demonstrated to have the capacity to predict characteristics of the experimentally synthesised structures of Phases A and F of the cobalt succinate materials, the arrangements of the structures produced indicate that, to realise a greater degree of precision, there is a requirement for the incorporation of moves that permit alteration of the internal coordinates of the entities involved in the MOF-synthesis simulations. Nevertheless, given that the structures formed in the MOF-synthesis simulation are influenced by the interactions between the reacting entities, the fact that the simulation method has the ability to correctly generate entities with the characteristics of the experimental structures demonstrates the capacity of the implicit-solvent method to accurately represent the interactions between the building blocks. Furthermore, that the extents to which the MOF-synthesis simulations progressed took only a small number of hours demonstrates the suitability of the implicit-solvent method for application in such simulations.

The execution of simulations carried out under the respective reaction conditions under which Phases A and F of the cobalt succinate materials can be synthesised experimentally required prior definition of a number of specifications. In particular, the ratio with which the different moves involved in the simulations were attempted, and the value of enthalpy associated with the simulated reactions were specified in relation to the requirement that the simulation adhered to the detailed balance condition. However, due to their associated computational costs, the simulations carried out as a means of representing the synthesis of the cobalt succinate materials did not reach equilibrium. Therefore, the suitability of these specifications could not be fully evaluated with respect to the precise considerations relevant to their definitions. Nevertheless, the fact that the simulations involving the specified parameters have the capacity to accurately predict characteristics of the experimentally synthesised cobalt succinate materials of interest is indicative of the appropriateness of the values utilised.

CHAPTER 6

Evaluation of the applicability of ideal adsorbed solution theory to mixture adsorption in MOFs

As expressed previously, MOF materials show promise for the adsorption-based separation and purification of mixtures of gases.[197] However, the design and development of an adsorption-based separation process requires knowledge of both single-component and mixture adsorption equilibria and, whilst single-component isotherms can be obtained relatively easily, accurate experimental measurement of mixture isotherms is time-consuming and relatively expensive.[198] Therefore, the ability to accurately predict adsorption isotherms for the components in a mixture affords great advantage. Ideal Adsorbed Solution Theory (IAST) is a technique by which multi-component adsorption isotherms can be derived solely from single-component adsorption isotherm data. The ability of IAST to predict mixture

adsorption is, however, dependent on the characteristics of the adsorbent and the behaviour of the mixture of adsorbates. Consequently, foretelling the ability of IAST to predict the adsorption of a given mixture in a particular adsorbent is difficult.[199] Therefore, given the number of possible MOF structures and the associated significant number of potential applications, insight into the ability of IAST to predict mixture adsorption in relation to the adsorbent structure and the constituents of the mixture is of value.

Assessment of the accuracy of IAST predictions of mixture adsorption requires comparison with measured mixture isotherms. Obtaining the very large amount of data needed from experimental measurements of mixture adsorption is not practicable. Therefore, both the single-component adsorption data, used as the inputs to IAST calculations, and the mixture data required to evaluate the performance of IAST, were obtained by means of a Monte Carlo simulation method that has been shown to give accurate results for adsorption in MOFs.[200] This approach was also taken by earlier, more limited, studies of the applicability of IAST to MOFs, which showed that, for the MOFs studied, mixture adsorption is reasonably well predicted by IAST. Yang and Zhong[201] found that for mixtures of methane and hydrogen, and methane and carbon dioxide in IRMOF-1 and Cu-BTC, IAST results closely matched mixture simulation results. These findings have been confirmed, in part, by an investigation by Babarao *et al.*[202] in which good agreement was obtained between IAST predictions and simulations for a binary mixture of CO_2 and CH_4 in IRMOF-1, and work by Keskin *et al.*,[203] which supported the finding that for mixtures of CH_4 and H_2 in Cu-BTC, IAST predictions were in agreement with simulation results. In addition, Bae *et al.*[204] obtained good agreement between IAST predictions and simulated isotherms for the separation of a binary mixture of CO_2/CH_4 in a mixed-linker MOF and Liu *et al.*[205] found that IAST predictions corresponded closely to simulation results for mixtures of methane and hydrogen in interpenetrated MOFs, in which a second framework coexists inside the first. By comparing, and achieving good agreement between, IAST predictions and simulated isotherms for different compositions of mixtures of carbon dioxide and nitrogen and mixtures of carbon

dioxide and oxygen in Cu-BTC, Yang *et al.*[206] concluded that IAST was applicable for the prediction of the adsorption behaviour for that particular system. Thus far, the only examples of disparity between IAST and simulations of mixture adsorption have been found by Yang and Zhong[201] for the mixture of carbon dioxide and hydrogen in IRMOF-1 and Cu-BTC.

Although the capacity for IAST to predict mixture adsorption in MOFs has been demonstrated, evidence of the applicability of IAST to MOFs exists for only a small number of MOF structures and for a limited range of mixtures of light gases, for which, because the different gas molecules are of similar sizes and the intermolecular interactions are of comparable strengths, ideal behaviour is expected and thus significant deviations from IAST are not expected. This is not a sufficient basis for the systematic application of IAST to predict mixture adsorption in MOFs. In this work, in order to develop an understanding of the limitations of IAST in relation to the adsorbate properties and the structural characteristics of the adsorbents, the applicability of IAST was evaluated, through comparison with simulations, for mixtures of adsorbate molecules of different sizes, asphericities and polarities in a number of structurally differing MOFs. There are a number of reasons for the occurrence of deviations from IAST, as outlined, alongside a delineation of the underlying principles of IAST, in Section 6.1.

6.1 Ideal Adsorbed Solution Theory

Derived by Myers and Prausnitz,[207] IAST is a method for predicting the adsorption equilibria for components in a mixture, using only single-component adsorption data at the same temperature and on the same adsorbent. IAST is based upon three assumptions: (i) that the same surface area is available to all adsorbates, (ii) that the adsorbent is inert, and (iii) that the multi-component mixture behaves as an ideal solution (such that the mean strength of interaction is equal between all molecules of the solution) at constant spreading pressure and temperature.

For a particular component, *i*, in an ideal adsorbed phase:

$$\bar{f}_i = f_i^{\,0}(\pi)x_i, \qquad \textbf{6.1}$$

where $\bar{f}_i$ is the fugacity of component *i* in the bulk gas phase, $f_i^{\,0}$ is the fugacity for component *i* at the solution temperature, *T*, and the spreading pressure, π, of the binary mixture when the adsorbed and bulk gas phases are in equilibrium, and x_i is the mole fraction of component *i* in the adsorbed phase. For each component, $f_i^{\,0}$ and π are related via the integrated Gibbs isotherm:

$$\frac{\pi A}{RT} = \int_0^{f_i^0} n_i d\ln f_i \qquad \textbf{6.2}$$

where n_i is the amount adsorbed of component *i* (obtained from the single-component adsorption isotherm) at pressure, *P*, which, in this work, was related to the component fugacity, f_i, using the Peng-Robinson equation of state.[208] *A* is the surface area of the adsorbent and *R* is the ideal gas constant. At equilibrium, each component has the same spreading pressure:

$$\pi = \pi_i = \pi_j \quad i \neq j\,. \qquad \textbf{6.3}$$

Since the sum of the mole fractions of all components in the adsorbed phase is equal to 1, Equation 6.1 can be rearranged to give the following identity:

$$\sum_i \frac{\bar{f}_i}{f_i^{\,0}(\pi)} - 1 = 0 \qquad \textbf{6.4}$$

which can be solved for a given bulk fugacity, $\bar{f}_i$, for each component to obtain the value of $f_i^{\,0}$ for each component. Thus, the adsorbed phase composition, x_i, can be

calculated from Equation 6.1 and the number of moles, n_i^0, of component *i* that would be adsorbed from the pure gas at $f_i^0(\pi)$ can be found from the single-component isotherm for each component. The total amount adsorbed of the mixture, n_t, is determined using the expression:

$$\frac{1}{n_t} = \sum \frac{x_i}{n_i^0} . \qquad \textbf{6.5}$$

The actual number of moles adsorbed of each component is then:

$$n_i = x_i n_t . \qquad \textbf{6.6}$$

To calculate the spreading pressure using Equation 6.2, the functional forms of the single-component isotherms must be specified by fitting a continuous function to a discrete set of adsorption data obtained over a finite range of pressures. Therefore, in this work, IAST was applied to single-component adsorption isotherms fitted using an isotherm equation proposed by Jensen and Seaton:[209]

$$n(P) = KP\left\{1+\left(\frac{KP}{\alpha(1+\kappa P)}\right)^c\right\}^{-1/c} \qquad \textbf{6.7}$$

where $n(P)$ is the adsorbed amount at pressure P, K is Henry's law constant, α is a measure of capacity, c determines the radius of curvature of the isotherm and κ accounts for the compressibility of the adsorbed phase. An inherent feature of IAST is that prediction of multi-component adsorption requires extrapolation of the single-component isotherms outside the range of pressures over which the data were collected and to which the form of the isotherm was fitted and specified. This is a consequence of the equality of component spreading pressures at equilibrium and because, as shown in Equation 6.2, the spreading pressure is related to the amount adsorbed, data for the less strongly adsorbed component are extrapolated to a greater

extent. Extrapolation of single-component adsorption data can introduce a degree of uncertainty in IAST predictions. Therefore, the use of an equation which accounts for the compressibility of the adsorbed phase and, in so doing, provides an accurate description of adsorption behaviour at high pressures is advantageous when assessing the applicability of IAST.

Departures from IAST may be attributed to the dissimilarity of the adsorbate molecules, and to the heterogeneity of the adsorbent.[210] In relation to differences in the adsorbates, disparity between the adsorbate-adsorbate interactions in the adsorbed solution will effect deviations from ideality unless the adsorbate-adsorbate interactions are small in comparison with the adsorbate-adsorbent interactions. In this case adsorption is governed by the interaction between the adsorbate molecules and the adsorbent and the non-idealities resulting from the interactions between adsorbate molecules will be inconsequential.[211] In MOFs, because of their open structure, the adsorbate-adsorbent interactions are often much weaker than in conventional adsorbents, such as zeolites or activated carbons. As a result, the adsorbate-adsorbate interactions have greater influence,[212] and so differences in the strengths of interactions between different adsorbate molecules in a mixture are likely to cause significant deviations from ideality.

Heterogeneities in the adsorbent structure can have considerable influence on the applicability of IAST. In conjunction with differences in the adsorbate molecules in a mixture, adsorbent surface heterogeneities at length-scales of the same order as the size of the adsorbate molecules can, as a result of competition between different adsorbates for adsorption sites, render particular sites of the adsorbent inaccessible to one component in a mixture. Therefore, the assumption that the same surface area is available to all adsorbate molecules might be invalid. In addition, larger scale energetic or structural adsorbent heterogeneities can affect the capability of IAST to predict mixture adsorption. IAST predictions are calculated on the basis that all of the information pertaining to the adsorption of a particular adsorbate in a given adsorbent can be derived solely from the single-component isotherm and the fact

that, in a heterogeneous adsorbent, the isotherm represents the aggregate of the contributions from regions of different energy or geometry is not taken into account. Consequently, where, due to adsorbent heterogeneities, different regions of an adsorbent have disparate adsorption selectivities for the components in a mixture, deviations from IAST can result.[213] In comparison with conventional adsorbents, as a consequence of the innumerable coordination geometries possible through judicious selection of metal vertices and organic linkers, MOFs present much more structural diversity and exhibit wide variation in pore topology. Therefore, given the extent to which the different pores in one particular MOF can vary in terms of shape and size, the effect of adsorbent heterogeneities on the applicability of IAST to the prediction of mixture adsorption in MOFs is likely to be significant. The effect of palpable heterogeneities in the adsorbent structure can be estimated by assuming local adsorption (i.e., adsorption at a particular site/pore) to be described by IAST and applying IAST separately to distinct sites so as to implement a form of Heterogeneous Ideal Adsorbed Solution Theory (HIAST).[210] Further details of the application of HIAST are given, following delineation of the simulation method through which the adsorption isotherms required for appraisement of IAST were obtained and specification of the adsorbent structures and the adsorbates involved in the mixtures for which the ability of IAST was assessed, in Section 6.2.

6.2 Simulation Method and Models

6.2.1 Simulation Method

The single-component and mixture adsorption isotherms necessary to evaluate the applicability of IAST to MOFs were obtained by means of Monte Carlo simulations in the grand canonical ensemble. Grand canonical Monte Carlo (GCMC) simulation is a method by which equilibrium adsorption in porous materials can be studied at the molecular level. The statistical-mechanical basis for the GCMC technique is outlined in detail elsewhere.[56, 57] As in experimental measurements of adsorption, GCMC simulations are performed such that the chemical potentials of the

adsorbates, μ, the volume of adsorbent, V, and the temperature, T, are held constant, while the total number of molecules, N, is allowed to fluctuate. To sample equilibrium configurations characteristic of the specified thermodynamic state, the GCMC method employs insertion and deletion moves, in addition to the translational and, where appropriate, rotational moves, for which the criterion governing the probability of acceptance is specified in Chapter 2. The respective probabilities of acceptance, p_{ins} and p_{del}, of the insertion and deletion moves are dependent on the number of molecules in the system, as shown in Equations 6.8 and 6.9.

$$p_{ins} = \min\left[1, \left(\frac{V}{\Lambda^3 (N+1)}\right) \exp\{\beta\left(\mu - \Delta U_{ins}\right)\}\right] \quad \textbf{6.8}$$

$$p_{del} = \min\left[1, \left(\frac{\Lambda^3 N}{V}\right) \exp\{-\beta\left(\mu + \Delta U_{del}\right)\}\right] \quad \textbf{6.9}$$

where

$$\Lambda = \sqrt{\beta h^2 / (2\pi m)} \quad \textbf{6.10}$$

is the thermal de Broglie wavelength, in which expression h represents Planck's constant and m is the adsorbate molecular mass. $\Delta U_{ins} = U(N + 1) - U(N)$ represents the energy change associated with the addition of one molecule at a random position and $\Delta U_{del} = U(N - 1) - U(N)$ is the energy difference resulting from the removal of a randomly selected molecule. As before, $\beta = 1/k_B T$, in which k_B represents the Boltzmann constant. In this work, the chemical potentials were related to the pressure and composition of the bulk gas using the Peng-Robinson equation of state[208] and GCMC simulations were executed using the multipurpose simulation code Music.[96] Typically, simulations were run so as to allow 20 000 Monte Carlo moves, on average, for each adsorbed molecule to ensure that equilibrium was reached. The first 40 % of the moves were used for equilibration and, as such, were not included in the calculation of isotherms. To obtain reliable statistics, the number

of unit cells involved in a particular simulation was selected such that the minimum number of molecules adsorbed at a given pressure value was 100. Further, in mixture simulations, sufficient unit cells were used so as to ensure that, on average, a minimum of 10 molecules of the less strongly adsorbed component were present in the simulation.

GCMC simulation is well established as a means of gaining insight in to adsorption in MOFs. The ability of GCMC to reproduce experimental adsorption behaviour in MOFs has been reviewed in detail elsewhere.[200] In general, GCMC simulations of adsorption in MOFs have been shown to be reasonably successful at replicating the behaviour of the equivalent experimental systems provided the potential models used to represent the system are accurate. In this work, where possible, the simulation parameters employed to represent adsorbent-adsorbate systems have been taken from previous simulation studies that have demonstrated the capacity to accurately predict the experimental behaviour of the systems of interest. Where GCMC simulations of adsorption in MOFs have been found to differ from the corresponding experiments, the discrepancies have often been attributed to the fact that the simulations involve a representation of the perfect crystal structure whereas the real material used in experimental adsorption measurements is likely to contain impurities and/or defects.[214]

In this work, for the purpose of evaluating the capacity of IAST to predict adsorption in MOFs, both single-component isotherms (from which IAST predictions were made) and mixture isotherms were determined via simulation. Therefore, the fact that the GCMC simulations carried out involve the perfect MOF structures does not impact on this evaluation of IAST. Utilisation of the perfect crystal structures as part of a simulation-based evaluation of the capacity of IAST to predict adsorption in MOFs means that this work does not constitute an exacting assessment of the ability of IAST with respect to experimentally measured adsorption isotherms. However, because GCMC simulations are generally considered to reasonably accurately represent adsorption behaviour in MOFs, the applicability of IAST in terms of

experimental measurements of adsorption is assumed to be comparable with the capabilities of IAST in relation to simulated adsorption isotherms in MOFs, as is discussed in this work.

In all simulations, the adsorbents were treated as rigid frameworks and modelled atomistically with the atomic positions fixed from crystallographic data. To minimize the computational requirements of the GCMC simulations, for each adsorbate species the potential energy between the adsorbate and adsorbent framework was pre-calculated on a three-dimensional grid with 0.2 Å spacing. During the simulations, the potential energy at any particular position in the adsorbent was calculated by interpolation using the grid. The van der Waals interactions between adsorbate molecules were calculated using the pair-wise additive Lennard-Jones (LJ) 12-6 potential. For simulations of a single-component, *A*, LJ potentials were cut off at $5\sigma_{AA}$, at which distance the interaction energy is negligible. Likewise, for simulations of mixtures of components *A* and *B*, interactions were cut off at $5\sigma_{AA}$ where $\sigma_{AA} > \sigma_{BB}$. All of the force fields and mixing rules used in this work have previously been successfully employed to model adsorption in MOFs, as detailed elsewhere.[200] The Lorentz-Berthelot mixing rules were applied to calculate mixed Lennard-Jones parameters. Lennard-Jones parameters for framework atoms were taken from the Universal Force Field (UFF).[215] For simulations involving molecules such as CO_2, which is quadrupolar, electrostatic effects were non-negligible and so Coulombic interactions between adsorbate and adsorbent were accounted for using the Ewald summation technique.[69] Since the applicability of IAST to adsorption in MOFs would not be expected to vary significantly with temperature, all simulations were performed at 300 K, which is within the range of temperatures relevant to many adsorptive separation processes. This simulation method was employed to obtain adsorption isotherms in a number of adsorbent materials, the structures of which are described in Section 6.2.2.

6.2.2 MOF Structures

In order to gain insight into the applicability of IAST to MOFs in general, several MOFs with widely differing structures and pore geometries were investigated. IRMOF-1[216] is a member of the class of isoreticular MOFs and consists of Zn_4O clusters connected by 1,4-benzenedicarboxylate (BDC) linkers to form a three-dimensional porous cubic framework. Linkers in the structure alternate between pointing inwards and outwards, conferring upon the structure two alternating inter-connected cavity types: smaller cavities of approximately 11.2 Å and larger cavities with dimensions of approximately 14.4 Å. A second MOF with a three-dimensional pore structure, Cu-BTC, which was first reported by Chui *et al.*,[217] contains copper ions, Cu^{2+}, as metal centers coordinated to benzene-1,3,5-tricarboxylate (BTC) linker molecules such that the resulting framework has three types of pore: two different main cavities of square cross-sections with dimensions of 10.5 Å and 12.2 Å respectively, and tetrahedral side pockets, approximately 5.0 Å in diameter. MIL-68(V),[218] which is composed of vanadium octahedron vertices coordinated by terephthalate linkers to form unconnected one-dimensional channels, exhibits two distinct pore types: large hexagonal and small triangular channels with pore openings, respectively, of approximately 13.6 Å and 4.8 Å. MIL-47(V)[219] is composed of the same vertices and organic linker molecules as MIL-68 but has one-dimensional channels which are rhombic in shape with a pore diameter of approximately 7.4 Å. Larger, one-dimensional rhombic channels with a cavity diameter of approximately 10.4 Å are present in MOF-69A,[220] which is composed of infinite Zn-O-C columns stacked in parallel and connected by biphenyl linkers. All pore sizes are as calculated using the method of Gelb and Gubbins,[221] which yields the diameter of the largest sphere that can fit into the cavity without overlapping the framework atoms. The structures of the MOFs studied are shown in Figure 6.1 and the Lennard-Jones parameters for the framework atoms are given in Table 6.1. The partial charges on the framework atoms in the MOF structures examined are displayed in Table 6.2. The applicability of IAST for the prediction of mixture adsorption in the aforementioned MOF structures was evaluated for a number of

mixtures of adsorbates. The details of the models by which the adsorbate molecules were represented in GCMC simulations of adsorption are described in Section 6.2.3.

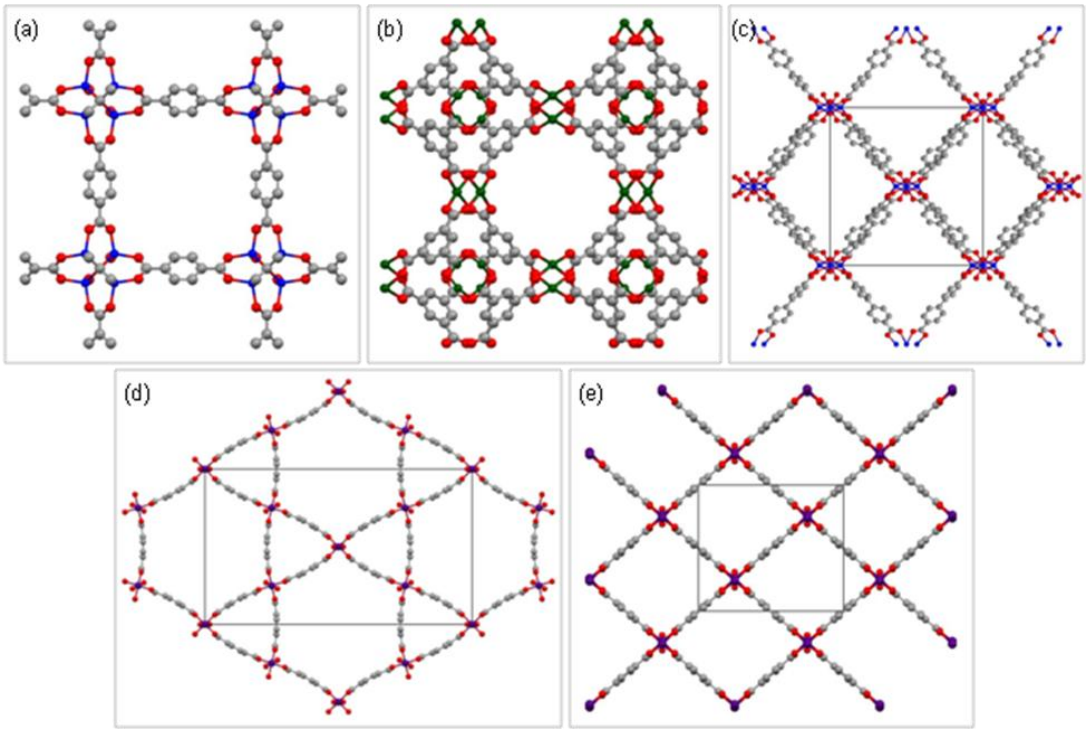

Figure 6.1: Crystal structures of the MOFs studied: (a) IRMOF-1; (b) Cu-BTC; (c) MOF-69A; (d) MIL-68; (e) MIL-47 where unit cells are delineated. (Zn, blue; Cu, green; V, purple; O, red; C, grey. H atoms are omitted for clarity.) Reproduced with permission from ref. [222]. Copyright 2012 American Chemical Society.

Table 6.1: Lennard-Jones parameters for atoms in the adsorbents studied, where ε is the LJ well depth, k_B is the Boltzmann constant, and σ is the atomic diameter. Parameters were taken from the Universal Force Field.[215] Reproduced with permission from ref. [222]. Copyright 2012 American Chemical Society.

Framework Atom	σ (Å)	ε/k_B (K)
H	2.571	22.14
C	3.431	52.8
O	3.118	30.19
Zn	2.462	62.4
Cu	3.114	2.516
V	2.8	8.05

Table 6.2; Partial charges on framework atoms in the MOFs studied.[223] Reproduced with permission from ref. [222]. Copyright 2012 American Chemical Society.

Framework Atom	IRMOF-1	Cu-BTC	MIL-68	MIL-47
C_1	0.76	-0.001	-0.071	-0.12
C_2	0.02	-0.131	0.598	0.76
C_3	-0.14	0.679	-0.067	-0.08
O_1	-0.76	-0.601	-0.4875	-0.52
O_2	-1.56	-	-0.613	-0.55
H	0.15	0.159	0.143	0.11
Zn	1.47	-	-	-
Cu	-	0.989	-	-
V	-	-	1.213	1.32

6.2.3 Adsorbate Models

For the range of MOF structures studied, the dependence of the applicability of IAST on the characteristics of the mixture of adsorbates was examined by comparing IAST predictions with GCMC results for mixtures of adsorbate molecules of differing sizes, asphericities and polarities. The applicability of IAST to mixtures of adsorbates of differing sizes was investigated by comparing IAST predictions with GCMC results for mixtures of adsorbates modelled as different sized spherical molecules using the united-atom approach. Thus H_2,[224] CH_4,[225] CF_4[226] and SF_6[226] were modelled as single spheres of increasing diameter.

By evaluating IAST predictions alongside GCMC results for mixtures containing short alkanes, the applicability of IAST for mixtures of adsorbates of differing degrees of asphericity was assessed. Using the united-atom approach, CH_3 and CH_2 groups were considered as single LJ spheres and ethane and propane were represented, respectively, as two and three uncharged interaction sites. The LJ

potential parameters were taken from the TraPPE force field developed by Potoff and Siepmann.[225]

Comparison of IAST predictions with simulation results for mixtures containing CO_2 allowed investigation of the applicability of IAST for mixtures of adsorbates of differing polarity. Carbon dioxide, which is a quadrupolar molecule, was modelled as a rigid linear tri-atomic molecule with a charged Lennard-Jones site located at each atom. The Lennard-Jones potential parameters for the atoms, for a CO_2 molecule with C-O bond length of 1.16 Å and bond angle $\angle$OCO of 180°, were taken from the TraPPE force field developed by Potoff and Siepmann.[227] The intrinsic quadrupole moment of CO_2 was represented by partial charges assigned to each atom. The charges for oxygen and carbon atom sites were, respectively, -0.35e and 0.7e. The Lennard-Jones parameters of all the adsorbates studied in this work are given in Table 6.3.

Table 6.3: Lennard-Jones parameters for the adsorbates used, where ε is the LJ well depth, k_B is the Boltzmann constant, and σ is the molecular diameter. Reproduced with permission from ref. [222]. Copyright 2012 American Chemical Society.

Adsorbate	σ (Å)	ε/k_B (K)	Ref.
H_2	2.958	36.7	[224]
CH_4	3.73	148.0	[225]
CF_4	4.662	134.0	[226]
SF_6	5.128	222.1	[226]
alkane CH_3 group	3.75	98.0	[225]
alkane CH_2 group	3.95	46.0	[225]
C (in CO_2)	2.8	27.0	[227]
O (in CO_2)	3.05	79.0	[227]

Of the structures described in Section 6.2.2, two MOFs, Cu-BTC and MIL-68, contain more than one type of cavity. As expressed in Section 6.1, the impact of such large-scale structural heterogeneities, which can affect the ability of IAST to predict the adsorption of mixtures of adsorbates, can be assessed by implementing a type of HIAST. The details of the simulation procedures involved in the application of the HIAST method employed in this work are described in Section 6.2.4.

6.2.4 HIAST Method

HIAST is an extension of IAST that enables consideration of the effect of palpable heterogeneities in the adsorbent structure.[210] Thus, HIAST can be applied to account for the intrinsic heterogeneities in the structures of MIL-68 and Cu-BTC, in which more than one kind of pore is present. Implementation of HIAST requires definition of individual sites which correspond to different pore types in the MOF. This can be achieved by creating, for a particular structure, a sitemap which takes the form of a three-dimensional grid with 0.1 Å spacing. Since different pore types represent distinct sites, by assigning each site a different number, the pores in the MOF can be allocated a site number. By defining mathematical inequalities relating to the boundaries of the pores, each grid point on the sitemap can be assigned a number corresponding to the site (e.g. type of pore) within which it is positioned. Thus, for MIL-68, where the smaller, triangular pores constitute one site and the larger, hexagonal pores represent a second site, using inequalities to distinguish between pores, a sitemap can be produced detailing the locations of the individual sites. The distinct adsorption sites accessible to a methane molecule, modelled as described in Section 6.2.3, in MIL-68 are illustrated in Figure 6.2. Sites in Cu-BTC can be defined using the same method, with the tetrahedral side pockets representing one site, and the two main cavities forming a second site.

When the adsorption simulation data are processed, the sitemap can be incorporated to obtain isotherms which describe the adsorption on separate sites. IAST predictions of adsorption, of a mixture of adsorbate species, for the individual sites

can be calculated from single-component GCMC data for adsorption of the species on the sites, which are identified as distinct types of pore in the MOF. HIAST predictions of mixture adsorption in a MOF structure represent the sum of these IAST predictions. Thus, HIAST constitutes a means of considering the effect of major structural heterogeneities on the ability of IAST to predict mixture adsorption. HIAST predictions determined, where appropriate to the system under investigation, via the procedure detailed are given in Section 6.3 alongside the results of the assessment of the applicability of IAST for the prediction of mixture adsorption in MOFs.

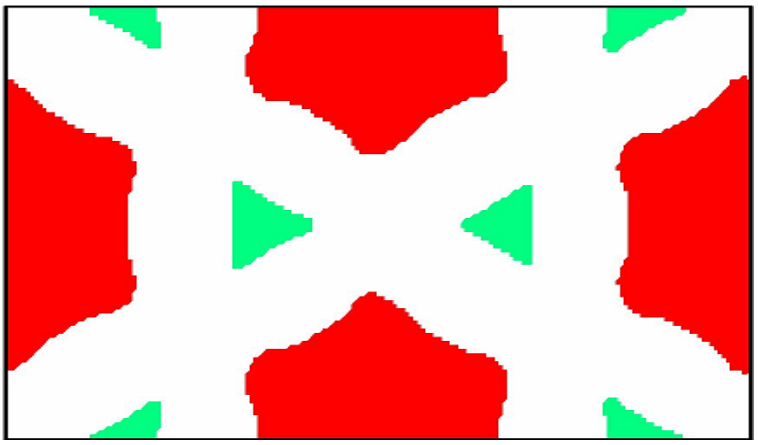

Figure 6.2: MIL-68 unit cell showing the adsorption sites available to a methane molecule where all areas coloured red represent one site (larger, hexagonal pores) and all areas coloured green represent a second site (smaller, triangular pores). The positions occupied by the framework atoms (i.e. where adsorption is not possible for methane) are coloured white. Reproduced with permission from ref. [222]. Copyright 2012 American Chemical Society.

6.3 Results

As expressed in Chapter 1, in adsorptive separation processes, the most sensitive measure of the separation capability is the selectivity of a porous material for different components in the mixtures. Therefore, the applicability of IAST was assessed by comparing values of selectivity calculated from IAST predictions alongside the equivalent selectivities obtained from GCMC simulation results. Evaluations of the ability of IAST to predict the adsorption in MOFs of mixtures of

adsorbates of differing sizes, asphericities and polarities are given, respectively, in Sections 6.3.1, 6.3.2 and 6.3.3.

6.3.1 Size of the Adsorbate Molecule

To investigate the applicability of IAST to systems with differing adsorbate sizes, IAST predictions were compared with simulation results for mixtures of the spherical adsorbates H_2, CH_4, CF_4 and SF_6. Because hydrogen is much more weakly adsorbed than the other adsorbates, single-component hydrogen isotherms were simulated and fitted for pressures up to 80 000 kPa. For equimolar mixtures of methane and hydrogen in the adsorbents IRMOF-1, Cu-BTC, MIL-68, MOF-69A and MIL-47 there are significant differences in the variation of the respective selectivities with pressure, as shown in Figure 6.3.

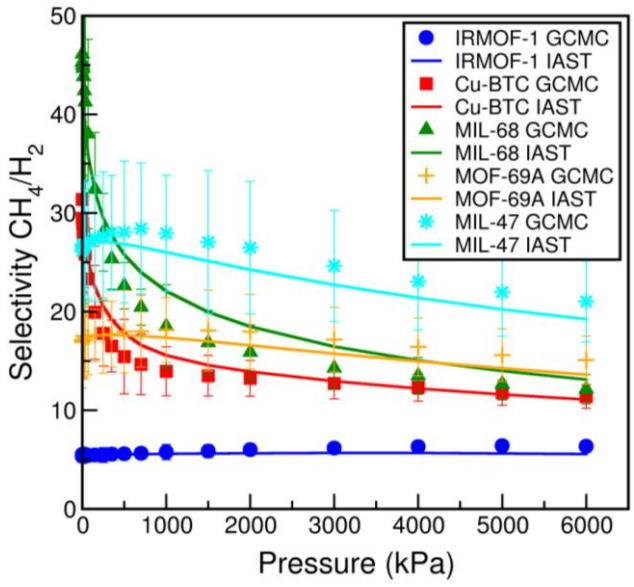

Figure 6.3: Selectivity for methane from an equimolar mixture of methane and hydrogen as a function of pressure in IRMOF-1, Cu-BTC, MIL-68, MOF-69A and MIL-47 at 300 K. Reproduced with permission from ref. [222]. Copyright 2012 American Chemical Society.

IRMOF-1, MOF-69A and MIL-47 exhibit virtually pressure-independent selectivity values, whilst the GCMC simulation results for MIL-68 and Cu-BTC display, with the initial pressure increase, a rapid decrease in selectivity, followed by more gradual

decline in selectivity at higher pressures. The characteristics of the adsorption behaviour are attributable to the differences in pore structure and topology of the MOFs. Whereas MOF-69A and MIL-47 have only one pore type and the pores in IRMOF-1 are similar in shape and dimension, both MIL-68 and Cu-BTC have two kinds of cavity of which one is smaller. Because the strength of adsorption is greater in the small pores, molecules are adsorbed preferentially in the smaller pores in MIL-68 and Cu-BTC. Thus, at lower pressures, the selectivity is high as result of the energetic advantage of the adsorption of the more strongly adsorbed component. Similarly, the higher selectivity for methane from an equimolar mixture of methane and hydrogen in MIL-47 as compared with MOF-69A reflects the greater strength of adsorption as a consequence of the smaller pores in MIL-47. The higher selectivity in MOF-69A than in IRMOF-1 results from the stronger interactions in the corners in MOF-69A. IAST predictions for the adsorption of the equimolar mixture of methane and hydrogen show good agreement with GCMC results with average relative errors between IAST and GCMC selectivities of 4 %, 4.7 % and 4.3 % for IRMOF-1, MOF-69A and MIL-47 respectively. The larger deviations between IAST and GCMC results evident at lower pressures in the adsorbents with more than one type of pore lead to marginally greater overall disparities of 6.2 % and 9.2 % between selectivity values calculated from IAST predictions and from GCMC results for, respectively, Cu-BTC and MIL-68.

To further investigate the influence of the inherent heterogeneities resulting from the two pore types in Cu-BTC (in which there are smaller, tetrahedral side pockets and larger, square pores) and MIL-68 (which contains smaller, triangular pores and larger, hexagonal pores), HIAST was also applied. Selectivity values calculated from HIAST are shown alongside IAST predictions and GCMC mixture simulation results for Cu-BTC and MIL-68 in Figure 6.4. For both structures HIAST results show good agreement with GCMC simulations, indicating that the deviations between IAST predictions and GCMC results are a consequence of the structures of the MOFs being heterogeneous.

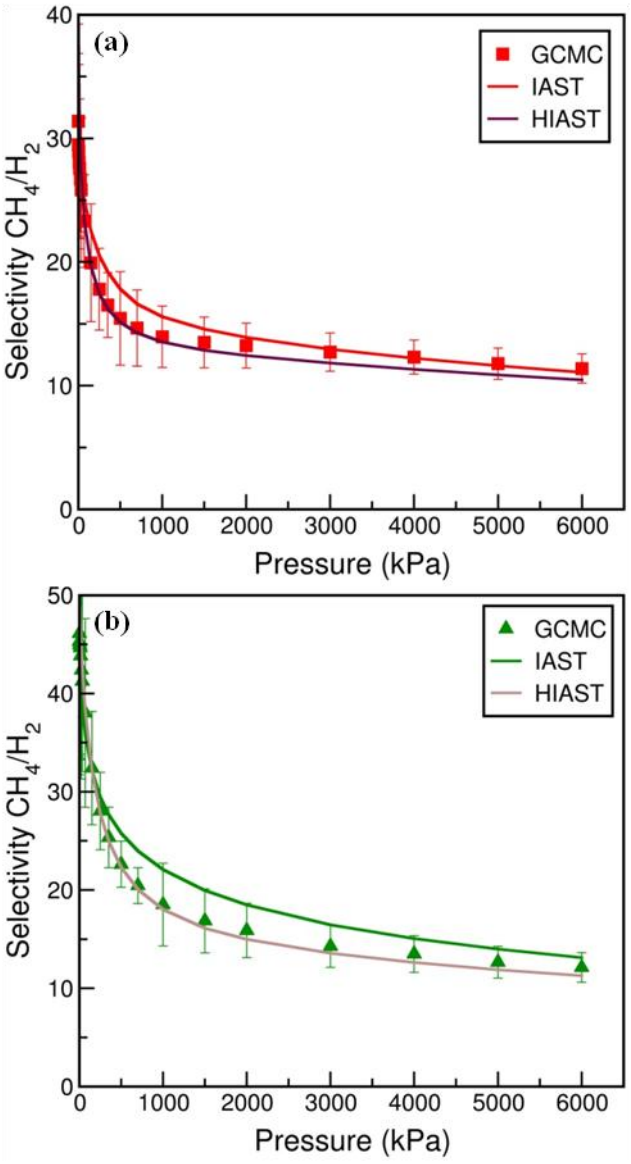

Figure 6.4: Selectivity for methane from an equimolar mixture of methane and hydrogen as a function of pressure at 300 K in (a) Cu-BTC, and (b) MIL-68. Adapted with permission from ref. [222]. Copyright 2012 American Chemical Society.

The effects of differences in adsorbate size on the applicability of IAST were investigated further through examination of a mixture of CF_4 and CH_4. The variation with pressure of the adsorption selectivity for CF_4 from an equimolar mixture of CF_4 and CH_4 is shown in Figure 6.5. The GCMC selectivities are generally well predicted by IAST throughout the range of pressures examined, with average relative errors between selectivity values calculated from IAST and GCMC results significantly below 10 % for all adsorbents studied with the exception of MIL-47, for

which large differences between GCMC and IAST in the low pressure region contributed to a higher average deviation of 12.6 %. This good agreement between IAST predictions and GCMC simulation data was not unexpected, given the similarity of the magnitudes of the Lennard-Jones parameters for CF_4 and CH_4, since the strengths of the interactions between adsorbate molecules are of similar magnitude for all molecules in the binary mixture, which therefore resembles an ideal solution. The good agreement for the mixture of CF_4 and CH_4 is in contrast to the results for the mixture of methane and hydrogen, for which the GCMC simulation results are much less well predicted by IAST in all adsorbents studied at low pressures. This indicates that the effects of the heterogeneities in MIL-68 and Cu-BTC are less significant for the adsorption of the mixture of CF_4 and CH_4.

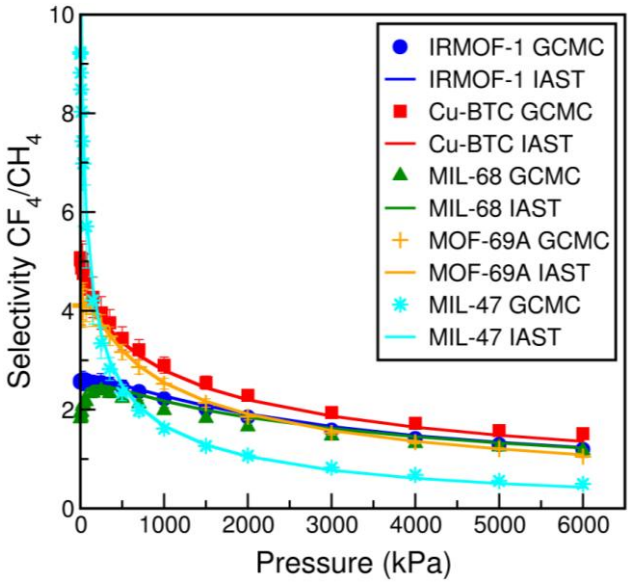

Figure 6.5: Selectivity for CF_4 from an equimolar mixture of CF_4 and CH_4 as a function of pressure in IRMOF-1, Cu-BTC, MIL-68, MOF-69A and MIL-47 at 300 K. Reproduced with permission from ref. [222]. Copyright 2012 American Chemical Society.

To investigate the differences in the influence of the adsorbent heterogeneity on the applicability of IAST for the CF_4/CH_4 mixture and for the CH_4/H_2 mixture, the selectivity values in the two distinct pore types in both MIL-68 and Cu-BTC were calculated separately for the two mixtures. As shown in Figure 6.6, in both Cu-BTC and MIL-68 there is great disparity between the selectivities in the different kinds of pore in a particular MOF for the mixture of methane and hydrogen, whilst the

selectivity values for adsorption of the mixture of CF_4 and CH_4 are of similar magnitude for the different pore types in a given structure.

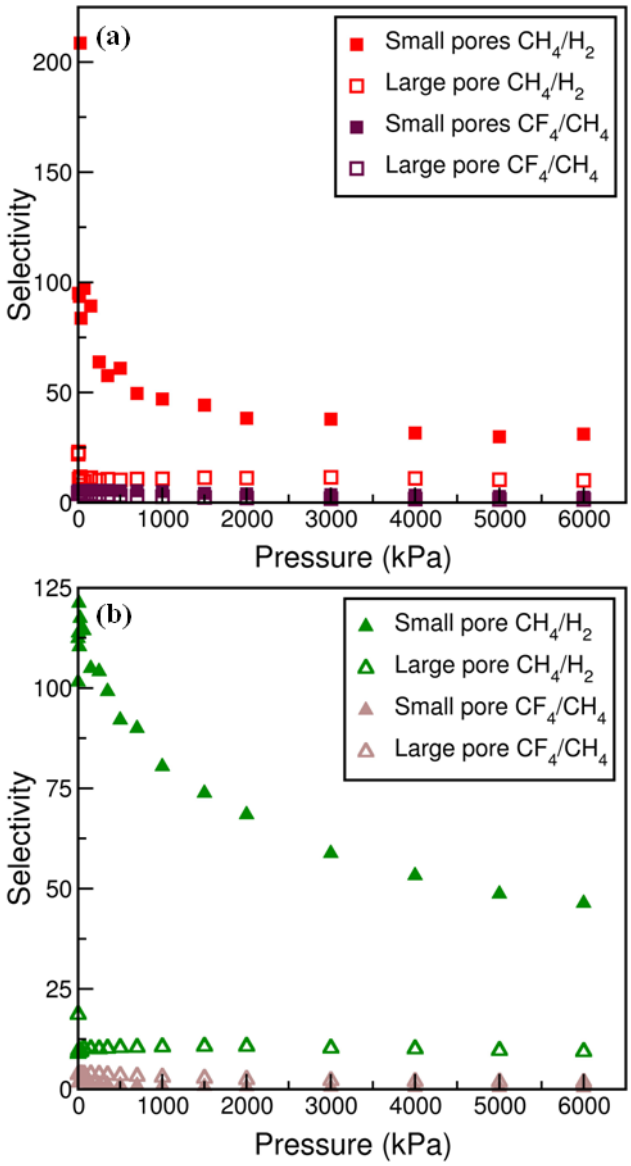

Figure 6.6: Comparison of adsorption selectivity values for CH_4 from an equimolar mixture of CH_4 and H_2 and for CF_4 from an equimolar mixture of CF_4 and CH_4 in the respective pores in (a) Cu-BTC, and (b) MIL-68, as calculated from GCMC simulations. Adapted with permission from ref. [222]. Copyright 2012 American Chemical Society.

Figure 6.6 also demonstrates that in comparison with the selectivities for methane from equimolar mixtures of methane and hydrogen, the selectivities in the smaller pores in both Cu-BTC and MIL-68 are much lower for the mixture of CF_4 and CH_4.

This is because there is much less disparity in the strengths of adsorption of CF_4 and CH_4. That the selectivity for CF_4 from an equimolar mixture of CF_4 and CH_4 is lower in the smaller pores than in the larger pores in the respective structures is a consequence of the limited size of the small pores such that there are few adsorption sites available to the more strongly adsorbed, but larger, CF_4 molecules. Thus, the heterogeneities of MOFs with different kinds of pores are influential as regards the applicability of IAST only for mixtures for which there is significant disparity between the adsorption selectivities in the distinct pore types within the structure.

The effects of greater disparity in adsorbate size were assessed by evaluating the applicability of IAST for the prediction of a mixture of SF_6 and CH_4. The variation with pressure of the adsorption selectivity for SF_6 from an equimolar mixture of SF_6 and CH_4 is illustrated separately for lower and higher pressure regions in Figure 6.7 in order to examine the applicability of IAST to the prediction of SF_6/CH_4 mixture adsorption throughout the pressure range studied without being subject to the influence of the magnitude of scale necessary to display the high selectivity values at low pressures. Given the relative strengths of adsorption of SF_6 and CH_4, the high selectivities for the adsorption of SF_6 over CH_4 exhibited by all adsorbents at low pressures, as shown in Figure 6.7a, were expected. Over the pressure range studied, in all adsorbents examined there were significant deviations between selectivity values calculated from IAST predictions and those from GCMC simulations. The average deviations between IAST and GCMC results were particularly high for MOF-69A and MIL-47, amounting to 42 % and 34 % respectively. Slightly lower were the average relative errors between selectivities calculated from IAST predictions and GCMC simulations for IRMOF-1, Cu-BTC and MIL-68 with, respectively, values of 26 %, 26 % and 21 %. The greater deviation, for all adsorbents, between IAST predictions and GCMC results for the SF_6/CH_4 mixture, in comparison with the average deviations for the CF_4/CH_4 and CH_4/H_2 is in agreement with the assertion of Dunne and Myers[228] and Sircar,[213] stating that IAST will fail if the co-adsorbates differ substantially in size, on the basis of studies performed in zeolite 13X and BPL-activated carbon adsorbents respectively.

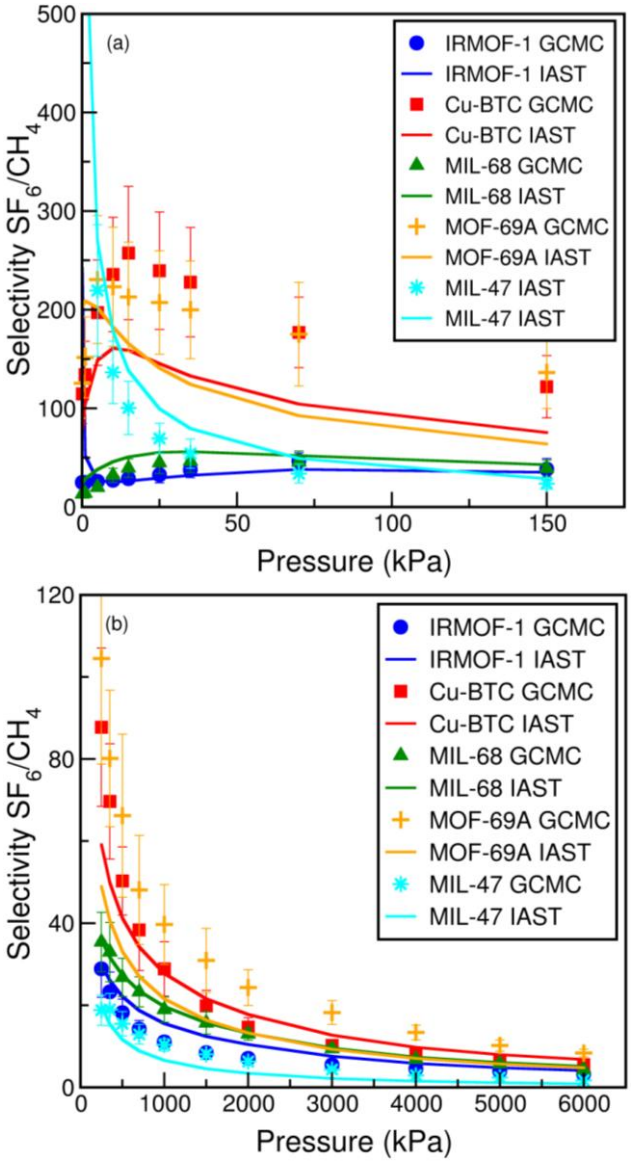

Figure 6.7: Selectivity for SF_6 from an equimolar mixture of SF_6 and CH_4 as a function of pressure in IRMOF-1, Cu-BTC, MIL-68, MOF-69A and MIL-47 at 300 K for (a) pressures up to 150 kPa and (b) pressures above 150 kPa. Reproduced with permission from ref. [222]. Copyright 2012 American Chemical Society.

The disparity between IAST predictions and GCMC simulation results evident for all adsorbents is due to competition for adsorption sites. The larger SF_6 molecules are much more strongly adsorbed and therefore block adsorption sites to the smaller, less strongly adsorbed methane molecules, with the effect that unequal surface areas are available to the different adsorbate molecules, therefore contravening one of the

main assumptions of IAST. The competition for adsorption sites by SF_6 appears particularly marked in MOF-69A, in which the higher energy regions of the pore are, by the nature of their positions in the acute angles of the rhombic pores, fairly easily obstructed by the adsorption of SF_6. Because the effect of this competition in MOF-69A, such that the methane molecules cannot access the higher energy regions in the pore, cannot be predicted by IAST, the adsorption of methane is significantly over-predicted by IAST. Thus, the deviation between IAST and GCMC selectivity values is much more significant for MOF-69A than MIL-68, in which, because adsorption is principally in the large hexagonal pores, there exist no such isolated high energy regions and no sites which can be so easily blocked.

In the evaluation of the effects of differing adsorbate sizes on the applicability of IAST for the prediction of mixture adsorption in MOFs, for larger adsorbate molecules standard GCMC simulations in Cu-BTC are inconsistent with the equivalent experiments.[229] In Cu-BTC, the small pockets are accessible from the main pores through triangular windows of diameter 4.6 Å, without taking into account the van der Waals diameters of the framework atoms.[229] Therefore, experimentally, the small pores in Cu-BTC are inaccessible to CF_4 and SF_6 molecules due to size restrictions. However, in standard GCMC simulations, molecule insertions are attempted in the whole volume of the MOF and the restrictive effects of the triangular windows are not imposed. Thus, to gain insight into the applicability of IAST to the prediction of the adsorption of the SF_6/CH_4 and SF_6/CF_4 mixtures by means of GCMC simulations consistent with experiments in Cu-BTC, the CF_4 and SF_6 single-component and GCMC simulations for both mixtures were repeated with the restriction that the adsorption of CF_4 and SF_6 molecules in the small pockets was not permitted.

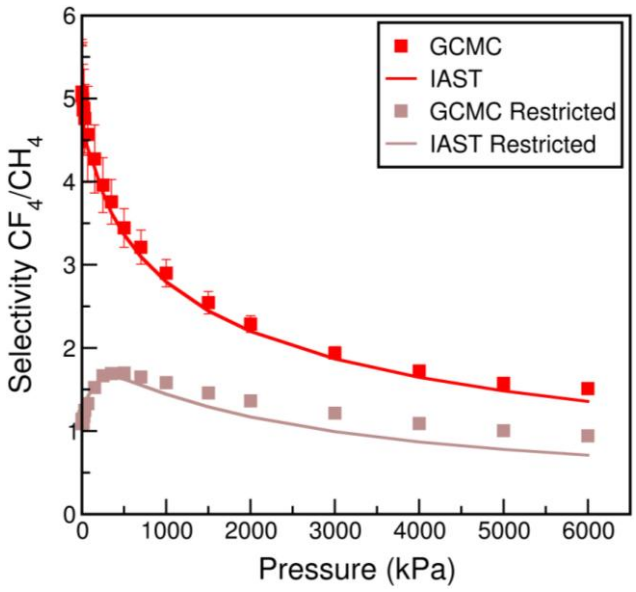

Figure 6.8: Selectivity for CF_4 from an equimolar mixture of CF_4 and CH_4 as a function of pressure in Cu-BTC at 300 K, where restricted results represent the selectivities where CF_4 has no access to the small pockets in Cu-BTC. Reproduced with permission from ref. [222]. Copyright 2012 American Chemical Society.

As is evident in Figure 6.8, when the adsorption of CF_4 is not permitted in the small pockets in Cu-BTC, the selectivity for CF_4 from an equimolar mixture of CF_4 and CH_4 is lower throughout the pressure range studied, most significantly at low pressures. This demonstrates the importance of considering only accessible pore space in GCMC simulations. The reduction in selectivity was expected since, by blocking the adsorption of CF_4 in the small pore, where adsorbate molecules are preferentially and more strongly adsorbed at low pressures, selective adsorption of CF_4 over CH_4 may take place only in the larger main cavity in Cu-BTC. Also apparent is that when the restriction of CF_4 adsorption is applied to simulations, the disparity between IAST predictions and GCMC results is higher, with an average relative error of 9.7 %, whereas the average deviation between IAST and GCMC selectivities for the unrestricted results is 5.2 %. Given that blocking the access of CF_4 to the small pore in Cu-BTC contravenes the inherent assumption of IAST that equal surface areas are available to both adsorbates, this poorer agreement was not unexpected. Nevertheless, even using single-component isotherms from GCMC simulations with CF_4 confined to experimentally accessible regions of Cu-BTC, IAST adequately predicts the adsorption of a mixture of CF_4 and CH_4. This

simulation result shows that IAST would be applicable to the corresponding real system.

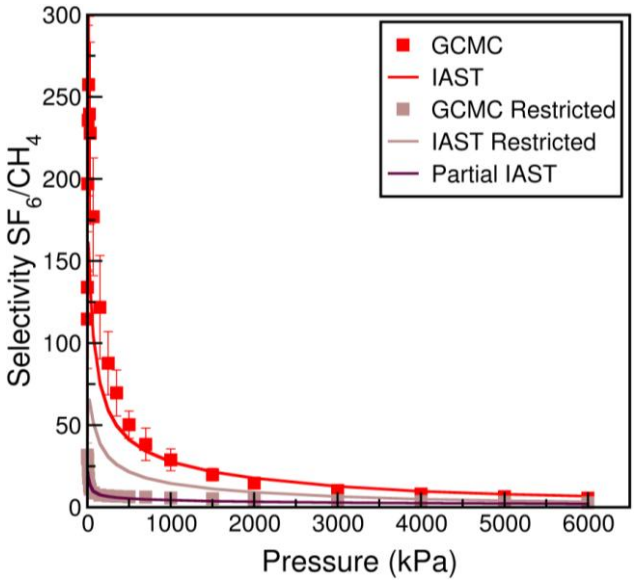

Figure 6.9: Selectivity for SF_6 from an equimolar mixture of SF_6 and CH_4 as a function of pressure in Cu-BTC at 300 K, where restricted results represent the selectivities where SF_6 has no access to the small pockets in Cu-BTC, and partial IAST results represent the sum of IAST predictions for the main cavities and single-component GCMC results for methane adsorption in the small pockets. Reproduced with permission from ref. [222]. Copyright 2012 American Chemical Society.

As shown in Figure 6.9, the application of restrictions to the pore space accessible to SF_6 results in much lower adsorption selectivities for SF_6 from an equimolar mixture of SF_6 and CH_4 than when no exclusions are in place. Again this is a consequence of selective adsorption in only the larger main cavity in Cu-BTC. The disparity between selectivities calculated from IAST predictions and those from GCMC simulations is much more significant for the mixture of SF_6 and CH_4 when the adsorption of SF_6 is confined to the main cavities in Cu-BTC. To minimize the effects of the inequality of the areas accessible to the different adsorbates, the different sites in Cu-BTC were treated individually such that results (denoted partial IAST) were obtained from IAST performed using the single-component isotherms obtained from GCMC simulations in the main cavities in Cu-BTC in addition to the single-component isotherm representing methane adsorption in the small pockets. Partial IAST results show much better agreement than restricted IAST results when

compared with the GCMC results obtained with a restriction imposed on the adsorption of SF_6. Thus, for the prediction of mixture adsorption in an adsorbent which presents restrictions to the adsorption of one component, by adhering as closely as possible to the underlying assumptions of the theory in the application of IAST, mixture adsorption can be well predicted. Owing to the greater difficulty in obtaining experimental isotherms for distinct regions of a structure, the application of this method to the equivalent real system would be much more complicated.

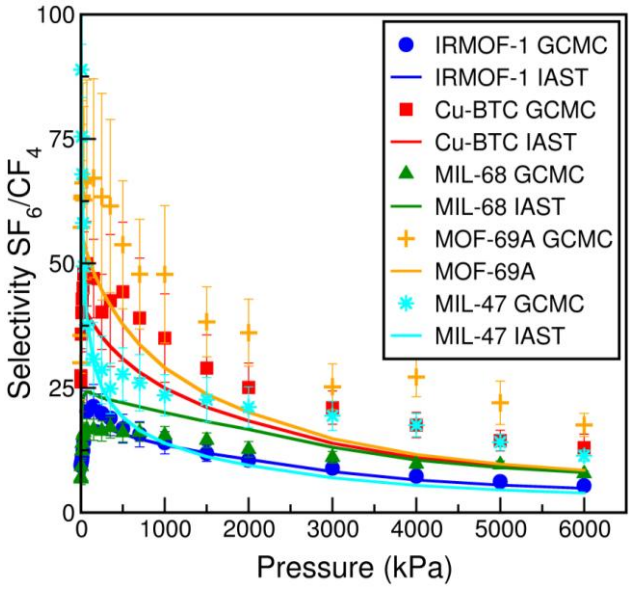

Figure 6.10: Selectivity for SF_6 from an equimolar mixture of SF_6 and CF_4 as a function of pressure in IRMOF-1, Cu-BTC, MIL-68, MOF-69A and MIL-47 at 300 K. Reproduced with permission from ref. [222]. Copyright 2012 American Chemical Society.

The effects of differences in adsorbate size on the applicability of IAST were further assessed by examining a mixture of SF_6 and CF_4. The variation in adsorption selectivity with pressure for SF_6 from an equimolar mixture of SF_6 and CF_4 is shown in Figure 6.10. As a consequence of the inability of a typical GCMC simulation to detect the inaccessibility of the small pockets in Cu-BTC to both SF_6 and CF_4 molecules, the single-component isotherms for SF_6 and CF_4, and the SF_6/CF_4 mixture adsorption isotherms for Cu-BTC were obtained with the condition that adsorption in the small pockets was not allowed. In general, there is better agreement between IAST predictions and GCMC results for the SF_6/CF_4 mixture than for the SF_6/CH_4

mixture. This is a consequence of the smaller difference in both adsorbate size and strength of adsorption for the molecules in the SF_6/CF_4 mixture in comparison with the SF_6/CH_4 mixture such that the competition for adsorption sites is less influential. The ability of IAST to predict mixture adsorption in MOFs was further assessed, through comparison with GCMC simulation results, for mixtures of adsorbate molecules differing in asphericity, as described in Section 6.3.2.

6.3.2 Asphericity of the Adsorbate Molecule

Evaluation of the applicability of IAST for mixtures of adsorbates of differing asphericities was achieved through comparison of IAST predictions and GCMC simulation results for mixtures of CH_4, C_2H_6 and C_3H_8. The variation with pressure of the adsorption selectivity for ethane from an equimolar mixture of ethane and methane is shown in Figure 6.11.

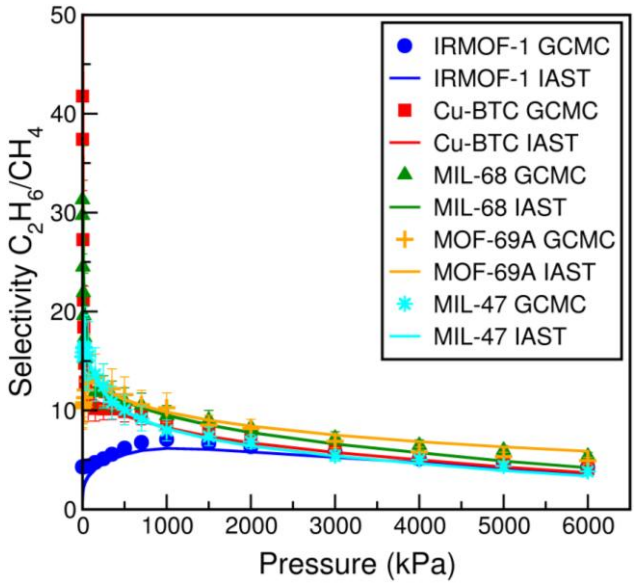

Figure 6.11: Selectivity for C_2H_6 from an equimolar mixture of C_2H_6 and CH_4 as a function of pressure in IRMOF-1, Cu-BTC, MIL-68, MOF-69A and MIL-47 at 300 K. Reproduced with permission from ref. [222]. Copyright 2012 American Chemical Society.

As shown in Figure 6.11, the selectivity values calculated from GCMC results are largely well predicted by IAST throughout the range of pressures examined, with average relative errors between IAST and GCMC results below 10 % for MIL-68,

MOF-69A and MIL-47. Despite the greater average discrepancies between IAST predictions and GCMC simulation results for Cu-BTC and IRMOF-1, which amounted to 13.8 % and 21 % respectively, there was generally good agreement throughout the pressure range studied. The agreement between IAST predictions and GCMC simulation results indicates that the differences in shape and adsorption strength of the different adsorbate molecules in a mixture of ethane and methane are not of sufficient magnitude as to effect deviations from IAST.

The effects of greater disparity in adsorbate asphericity were assessed by evaluating the applicability of IAST to the prediction of adsorption of an equimolar mixture of propane and methane. The variation with pressure of the adsorption selectivity for propane from an equimolar mixture of propane and methane is shown in Figure 6.12, in which the comparison between IAST and GCMC selectivity values is given for pressures below that at which the bulk mixture of propane and methane becomes a liquid. For the equimolar mixture of propane and methane, although IAST predictions deviate, quantitatively, from GCMC results, there is reasonable qualitative agreement. The largest disparities between IAST predictions and simulation results were obtained for those MOFs, Cu-BTC and MIL-68, in which there was greatest variation in selectivity with pressure, with average deviations of 33 % and 23 % respectively. This is a consequence of the inability of IAST to predict mixture adsorption in adsorbents where, due to heterogeneities, different regions of the adsorbent have greatly disparate selectivities. For IRMOF-1, MOF-69A and MIL-47, the average disparities between IAST and GCMC selectivity values were of similar magnitude, amounting to 14.5 %, 16.7 %, and 13.3 % respectively. Together with short-scale heterogeneities in the adsorbent structures, the considerable differences in both the sizes and adsorption strengths of the propane and methane molecules contribute to the deviation between IAST predictions and GCMC selectivity values since the ensuing competition for adsorption sites results in unequal areas available for adsorption of the different adsorbate molecules in the mixture.

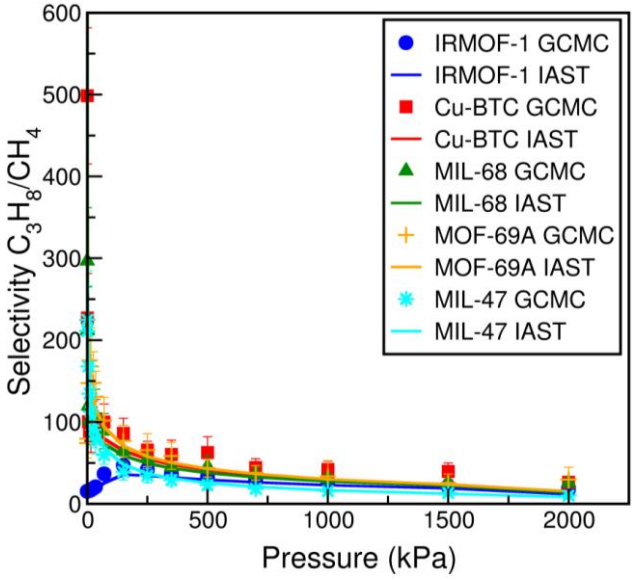

Figure 6.12: Selectivity for C_3H_8 from an equimolar mixture of C_3H_8 and CH_4 as a function of pressure in IRMOF-1, Cu-BTC, MIL-68, MOF-69A and MIL-47 at 300 K. Reproduced with permission from ref. [222]. Copyright 2012 American Chemical Society.

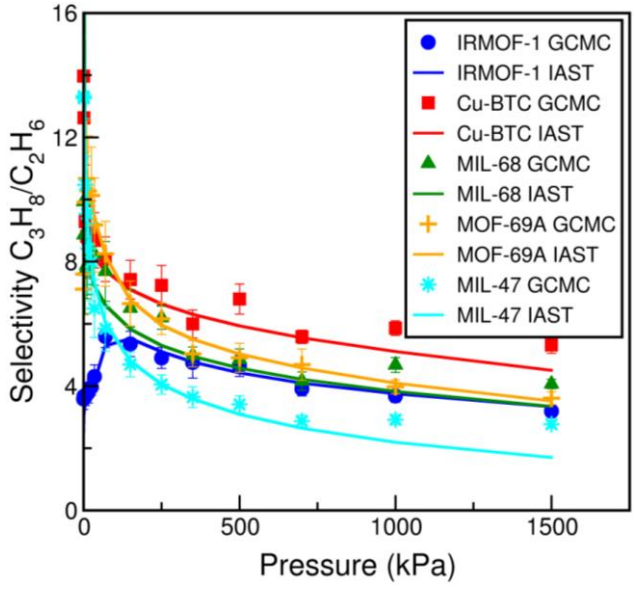

Figure 6.13: Selectivity for C_3H_8 from an equimolar mixture of C_3H_8 and C_2H_6 as a function of pressure in IRMOF-1, Cu-BTC, MIL-68, MOF-69A and MIL-47 at 300 K. Reproduced with permission from ref. [222]. Copyright 2012 American Chemical Society.

The effect of differences in adsorbate asphericity on the applicability of IAST was further assessed by examining a mixture of propane and ethane. The variation with pressure of adsorption selectivity for propane from an equimolar mixture of propane and ethane at 300 K is shown in Figure 6.13. As with the propane/methane mixture,

the comparison between GCMC and IAST results for the equimolar mixture of propane and ethane is shown only for pressures lower than that at which the bulk mixture becomes a liquid. In terms of the applicability of IAST to the prediction of mixture adsorption of an equimolar mixture of propane and ethane, the average deviation between IAST and GCMC results was lower than the corresponding value for the propane/methane mixture. However, for MIL-68, the average relative error between IAST predictions and GCMC simulation results of 22 % demonstrates only a marginal improvement in agreement for the mixture of propane and ethane. For IRMOF-1, Cu-BTC, MOF-69A, and MIL-47, the average discrepancies between selectivity values obtained from IAST and GCMC results were 6 %, 8.2 %, 10 % and 10.3 %, respectively. Thus, with lower disparity between both the adsorbate asphericities and the strengths of adsorption of the components in the mixture, the capacity for IAST to predict mixture adsorption is greater. The suitability of IAST as a means of predicting mixture adsorption in MOFs was further evaluated for mixtures of adsorbates of different polarities, as detailed in Section 6.3.3.

6.3.3 Polarity of the Adsorbate Molecule

The applicability of IAST to mixtures containing molecules of different polarity was assessed for equimolar mixtures of carbon dioxide and hydrogen. Because hydrogen is much more weakly adsorbed than carbon dioxide, IAST results were calculated using the functional forms of isotherms fitted to single-component hydrogen data simulated at pressures up to 80 000 kPa. As shown in Figure 6.14a, there is significant deviation between IAST predictions and simulated results for mixtures of CO_2 and H_2 in all adsorbents examined. Both MIL-68 and Cu-BTC display an initial decrease in selectivity with pressure, which reflects the presence of distinct small pores in their structures. The average relative error between IAST and GCMC selectivity values was significant for all adsorbents studied, with the most sizeable average deviation, with a value of 67 %, between IAST and GCMC results for IRMOF-1. Average deviations between IAST predictions and simulation results for

Cu-BTC, MIL-68 and MIL-47 were considerably lower, amounting to 30 %, 30 % and 40 % respectively.

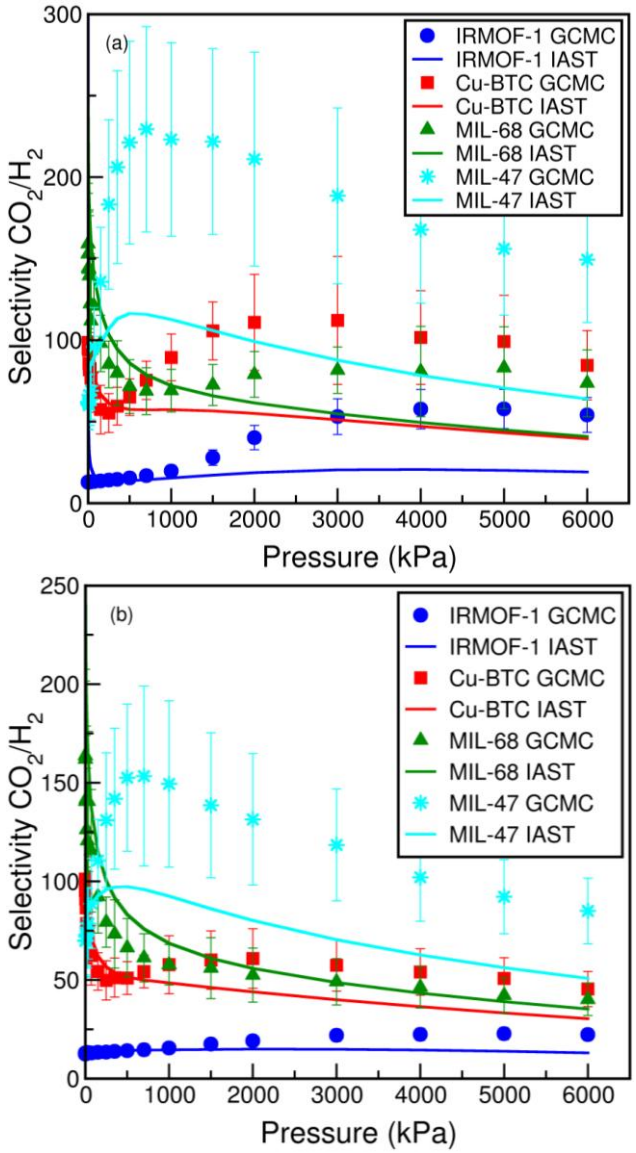

Figure 6.14: Selectivity for CO_2 from an equimolar mixture of CO_2 and H_2 as a function of pressure in IRMOF-1, Cu-BTC, MIL-68 and MIL-47 at 300 K with (a) CO_2-CO_2 and CO_2-framework electrostatic interactions included and (b) CO_2-CO_2 electrostatic interactions excluded. Reproduced with permission from ref. [222]. Copyright 2012 American Chemical Society.

For the equimolar mixture of carbon dioxide and hydrogen the failure of IAST to predict the mixture adsorption behaviour is most pronounced at higher pressures, indicating that the poor agreement is a consequence of the non-ideality of the

adsorbed phase since, at higher pressures, with greater loading of each adsorbate, the adsorbed phase less closely resembles an ideal solution. Due to the differing polarities of CO_2 and H_2, the mixture deviates significantly from ideality: as a result of the electrostatic interactions of the quadrupolar CO_2 molecules, there is great disparity in the strengths of interaction between different components in the mixture. In order to elucidate the effect of the electrostatic interactions between CO_2 molecules on the applicability of IAST to the prediction of the adsorption of a mixture of CO_2 and H_2, the simulations were repeated without taking into account the intermolecular CO_2 electrostatic interactions, while calculating the CO_2-framework interactions as before. As shown in Figure 6.14b, with the exclusion of electrostatic interactions between CO_2 molecules, there is much better agreement between selectivity values calculated from IAST predictions and from GCMC simulations. Quantitatively, the average relative errors between IAST predictions and GCMC simulation results were lower, for all adsorbents studied, than when CO_2-CO_2 electrostatic interactions were included. Average discrepancies between IAST and GCMC results obtained when intermolecular CO_2 electrostatic interactions were not taken into account amounted to 14.2 %, 18 %, 23 %, and 26 % for IRMOF-1, CU-BTC, MIL-68 and MIL-47 respectively. Thus, by disregarding the intermolecular CO_2 electrostatic interactions, and consequently lessening the disparity in the strengths of interactions between different constituents of the mixture, a considerable improvement in the agreement between IAST predictions and GCMC results was found. This confirms that the failure of IAST to predict the CO_2/H_2 mixture adsorption was partially due to the non-ideality of the adsorbed phase, resulting from the differing polarities of the components in the mixture.

That the deviation between IAST predictions and GCMC simulation results is significant even in the absence of intermolecular CO_2 electrostatic interactions shows that the poor agreement cannot be attributed solely to the non-ideality of the adsorbed phase. Given the substantial inequalities in terms of the shape and strength of adsorption of CO_2 and H_2, the disagreement between IAST predictions and GCMC simulation results is likely to result from competition for adsorption sites

causing disparity in the areas available for adsorption of different adsorbate molecules. This effect is particularly marked in MIL-47 as the small pore size restricts the number of adsorption sites.

6.4 Conclusions

In this chapter, the ability of IAST to predict binary mixture adsorption in MOFs was assessed through comparison with GCMC simulations. By means of examination of the predictive capability of IAST for binary mixtures of adsorbates of differing sizes, asphericities and polarities in a range of MOF structures, the applicability of IAST was studied in relation to both the properties of the mixtures and the structural features of the MOFs.

In general, in MOFs that present no major heterogeneities, such as cavities of very different sizes, the adsorption of mixtures of adsorbates of similar sizes and strengths of adsorption can be well predicted by IAST, such that average deviations between IAST predictions and GCMC mixture simulation results are typically below 10 %. However, ranges of approximately 20 – 40 % and 15 – 35 % for the average deviations between IAST and GCMC results for, respectively, the SF_6/CH_4 and C_3H_8/CH_4 mixtures in the MOFs studied showed that IAST is less able to predict the adsorption of mixtures of adsorbates of disparate sizes or asphericities. For such mixtures, in conjunction with the adsorbate-scale heterogeneity in the cavities presented for adsorption in MOFs, the competition between different adsorbates for adsorption sites results in situations which deviate from the underlying assumptions of IAST.

Particularly significant deviations from IAST have been found to arise from adsorbed phase non-idealities. Where, as a consequence of the significant disparity in the strengths of interaction between molecules in a mixture of adsorbates of differing polarities, the adsorbed phase does not resemble an ideal solution, the ability of IAST to predict mixture adsorption is particularly poor. This was evident for the mixture

of CO_2 and H_2, for which average deviations between IAST and GCMC results ranged from 30 % to 67 %.

CHAPTER 7

Conclusions and further work

As the culmination of this thesis, this chapter provides a discussion of the conclusions and findings of the work described in the previous chapters. In addition, this chapter supplies suggestions for areas in which additional work would be worthwhile. The conclusions arising from the work undertaken in this thesis are presented in Section 7.1, and recommendations for further work are described in Section 7.2.

7.1 Conclusions

The main body of work in this thesis was undertaken in pursuance of the aim of developing a simulation method with the capacity to represent the self-assembly process involved in the synthesis of MOFs to the extent of being able to accurately predict the MOF structures that form under specified reaction conditions. With this objective, as described in Chapter 2, a canonical Monte Carlo simulation,

incorporating an explicit-solvent representation, of the self-assembly of the vertices (cobalt ions) and linkers (succinate ions) associated with the formation of the series of cobalt succinate materials was carried out. In this explicit-solvent simulation the self-assembling vertex and linker components were modelled almost atomistically using established molecular potentials. The use of an explicit-solvent approach was found to be impracticable, in terms of the associated computational demands, for the purpose of simulating the synthesis of MOFs. This instigated the development of an implicit-solvent approach, which is outlined in Chapter 3. This approach involves, as a means of representing the effective interactions between solute species, the utilisation of PMFs determined between all pairs of the atom constituents of the initial reactants in the cobalt succinate system. These PMFs were calculated via umbrella sampling followed by WHAM analysis. In the umbrella sampling simulations, the atoms were modelled using the parameters with which they were represented in the respective initial reactant ions. The PMF-calculation approach utilised was found to yield PMF profiles with features that relate to the physical behaviour of the species in question.

The ability of the implicit-solvent approach to replicate the equivalent explicit-solvent behaviour was assessed for all possible pairs of the initial reactants that are involved in the formation of the cobalt succinate materials, as described in Chapter 4. The implicit-solvent and explicit-solvent approaches were compared by evaluating the capacity of the implicit-solvent method to reproduce the representative configurational sampling resulting from the use of the explicit-solvent representation. In order for the explicit-solvent sampling simulations to be computationally feasible, comparison of the implicit-solvent and explicit-solvent representations, by necessity, involved the use of constraining potentials to restrict the sampling close to a specified ion-ion separation distance. The implicit-solvent approach was demonstrated to require the employment of a hard-sphere potential in place of the PMF between the cobalt ion and the carbon atom, which exists largely surrounded by other species, in the succinate ion. The implicit-solvent and explicit-solvent representations were found to agree acceptably well at the separation distances that are most important when the respective ions interact in simulations of the synthesis

of MOFs. In comparing the implicit-solvent and explicit-solvent approaches the far lower computational demands of the implicit-solvent method were demonstrated. Thus, the implicit-solvent approach was concluded to have the capacity to enable the execution of simulations in which solvent molecules are not explicitly included, and was therefore considered suitable for application in simulations of the synthesis of MOFs.

To represent the synthesis of MOFs, a simulation method involving a kinetic Monte Carlo approach in which the physical interactions between reacting entities are considered to be more important than the energetics of the actual reactions between the entities was developed, as described in Chapter 5. This simulation method was applied under the different reaction conditions corresponding to the experimental syntheses of Phase A and Phase F of the cobalt succinate materials. That the two studied cobalt succinate phases are produced under different reaction conditions enabled evaluation of the capacity of the simulation method to predict the structure formed in relation to specified synthesis conditions. The simulation method is limited, due to the associated computational demands resulting from the manner of implementation, in terms of the extent of the synthesis process that can be represented. Nevertheless, the MOF-synthesis simulation method was found to have the capacity to accurately predict characteristics of the experimentally synthesised structures of Phases A and F of the cobalt succinate materials when the different synthesis conditions under which the structures form experimentally were applied. Thus, the simulation method developed in this work was concluded to have some capacity to represent the synthesis of MOFs. However, the requirement for, in order to enable a greater degree of precision as regards the arrangements of the structures generated, the incorporation of moves that allow intramolecular alterations of the simulated species was demonstrated.

The ability of the MOF-synthesis simulation method to predict details of the MOF structures that are synthesised under particular reaction conditions confirmed the capacity of the implicit-solvent simulation method to realistically represent the interactions between the building blocks. In addition, the relatively short time scales,

of a matter of hours, associated with the extents of the MOF-synthesis processes simulated provided further evidence of the suitability of the implicit-solvent approach utilised. As the test system for the development of the implicit-solvent approach and the application of the MOF-synthesis method, the cobalt succinate system proved to be a good selection. The simplicity of this system, such that only two solute species were involved, and the fact that the solute species were composed of small numbers of constituents meant that the determination of the PMFs associated with the implicit-solvent method was fairly straightforward. Furthermore, the clearly distinct differences in the cobalt succinate phases that form under different reaction conditions allowed judgement of the abilities of the formulated simulation method as a means of representing the synthesis of MOFs.

As a means of crystal structure prediction, the MOF-synthesis simulation method demonstrates potential. By allowing the generation of candidate structures within an environment representative of that in which the experimental structures are synthesised, the MOF-synthesis method developed in this work has the potential to accurately predict the formation of different crystal structures under different reaction conditions. Many of the crystal structure prediction methods in existence are devoid of this capacity. The fact that the simulation method formulated as described herein can correctly predict characteristics of experimental structures that are synthesised from the same building blocks influenced by different reaction conditions is indicative of the merit of adopting this approach. Nevertheless, it is clear that, at present, the MOF-synthesis method developed in this work cannot predict crystal structures at the precise level of detail that has been achieved through the application of other prediction methods in studies examining different structures. However, with further development of the MOF-synthesis simulation described in this work, a greater level of exactness in terms of the structure-predicting capacity might be possible.

In Chapter 6, away from the main objective of the work, a systematic study of the applicability of IAST to MOFs was performed by using GCMC simulations to investigate the suitability of IAST for the prediction of adsorption of mixtures of

molecules of differing sizes, asphericities and polarities in a range of structurally differing MOFs. IAST was shown to be generally accurate for the prediction of mixture adsorption in MOFs. Where IAST was found to be less accurate, deviations resulted from both mixture effects, in the form of non-idealities in the adsorbed phase, and characteristics of the adsorbent structures. In terms of the MOF structure, departures from IAST were found to be a consequence of heterogeneities both on the scale of the unit cell and on shorter length-scales, whereby competition for adsorption sites has a strong influence. The HIAST method was implemented as a potential means of accounting for possible limitations of IAST that result from heterogeneities in the structures of MOFs. The HIAST method was found to be capable of accounting for large-scale heterogeneities in the structures of MOFs, and, therefore, represents a technique suitable for predicting mixture adsorption in structurally heterogeneous MOFs. Furthermore, in addition to the conclusions drawn from the results of examining the ability of IAST to predict mixture adsorption in MOFs, the results of the work undertaken in the assessment of the applicability of IAST serve to demonstrate the capabilities of MOFs as adsorbent materials suitable for separation processes.

7.2 Future Work

Advancement of the work in this thesis would be worthwhile in relation to two principal matters: the implicit-solvent method and the simulation method employed as a means of representing the synthesis of MOFs. Recommendations for further work in these areas are discussed in Sections 7.2.1 and 7.2.2, respectively.

7.2.1 Implicit-Solvent Method

Whilst the capabilities of the developed implicit-solvent method have been established, there are a number of areas with respect to which further exploration would be beneficial. In this work, to ensure the utmost precision, the PMFs were very painstakingly determined. This involved performing the appurtenant window simulations individually and for a high number of MC steps. The precision of this

approach is indicated by the negligibility of the errors associated with the PMF data. Nevertheless, obtaining the PMFs in this manner is fairly computationally expensive and there is merit in the consideration of means to lessen the demands of the task. This could be achieved by reducing the number of steps involved in the window simulations from which the PMFs are derived, or by continuous simulation of successive windows along the reaction coordinate. The impact, on the capabilities of the implicit-solvent method, of such changes to the PMF-calculation method could be assessed through comparison with the previously obtained evaluations of the behaviour of the solute ion pairs.

Furthermore, as regards the requirements associated with the implementation of the implicit-solvent method, consideration of the structural basis for the determination of PMFs is worthwhile. In this work, the implicit-solvent method is based on the calculation of PMFs between all of the atoms in the solute species. Because the methylene groups in the succinate ion are represented by a united atom, between the cobalt and succinate ions associated with the cobalt succinate system there are four different types of constituent entities, requiring the calculation of ten PMFs. However, if the MOF-synthesis simulation method and, by association, the implicit-solvent approach were applied to MOF systems involving species that are composed of greater numbers of atoms, the calculation of the PMFs between all pairs of atoms could be a significant task. Therefore, in such a situation, there would be value in considering whether such an approach is necessary for the purpose of representing the synthesis of MOFs. In this work, the level of solvent-structuring detail provided by a PMF was not suitable for the carbon-cobalt interaction as a result of the position of the carbon atom, which is largely surrounded by other atoms within the succinate ion. This suggests that an implicit-solvent approach based on the calculation of PMFs between all possible atom pairs is not entirely essential and that there might be some potential for reducing the number of PMFs required. Therefore, before applying the MOF-synthesis simulation method to systems containing different solute species, there would be some value in investigating, by examining different MOF systems, whether there is any possibility of identifying the characteristics of

atoms for which the related interactions need not necessarily be considered by means of PMFs.

7.2.2 MOF-Synthesis Simulation Method

Clearly, as regards the MOF-synthesis simulation method, the initial focus of further undertakings should be on the formulation of a means of implementing the simulation method such that the associated computational requirements are not restrictive. Ideally, the computational demands of the simulation method should present no barrier to the simulated system reaching equilibrium. This would allow a more thorough evaluation of the capacity of the simulation method to represent the synthesis of MOF materials and give greater insight as to the ability of the method to predict the MOF structures that form under specified reaction conditions. Furthermore, the development of the simulation method to enable MOF-synthesis simulations to proceed to equilibrium would allow examination of the suitability of parameters, such as the enthalpy change associated with the reaction moves and the ratio with which different move types are attempted. In this work, for the cobalt succinate system, these parameters were specified in relation to the requirement that, at equilibrium, the simulated system satisfied the detailed balance condition. Therefore, assessment of the simulated cobalt succinate systems at equilibrium would yield insight as to the appropriateness of the bases on which these parameters were established. This would provide valuable knowledge of the considerations relevant to the specification of parameters that must be defined prior to commencing the simulations of the synthesis of MOFs. Undoubtedly, greater understanding of the issues involved in the definition of simulation parameters is of vital importance for further application of the simulation method to other MOF systems

Beyond the establishment of a less computationally costly MOF-synthesis simulation method, the results in Chapter 5 indicated the requirement for the incorporation of moves that permit modification of the internal coordinates of the entities involved in the simulations of MOF synthesis. The introduction of such moves would require the formulation of accurate intramolecular potentials for both the initial reactants and

the species generated in reactions. An essential requirement of any molecular model is that the conformational properties of the species in question can be correctly reproduced. However, the determination of potentials that exactly represent all possible bond-stretching, bond-bending and torsional intramolecular movements for all entities involved in a MOF-synthesis simulation would be an enormous task. Therefore, there would be merit in investigating the extent of the intramolecular variation, in terms of both the relevant bonds and the manner of the intramolecular alteration, that is actually necessary for the simulation method to have the capacity to accurately predict structures corresponding to the experimentally derived structures. This will, of course, depend to some degree on the characteristics of the MOF structure under examination.

In employing the simulation method as a means of representing the synthesis of MOFs, whilst further assessment of the capabilities of the simulation method would naturally accompany any efforts to rectify the outlined deficiencies of the present formulation, there are a number of areas in which evaluation of the method would be pertinent. Examination of any impact, on the ability of the method to represent the synthesis of MOFs, of the size of the system and of the shape of the simulation cell would yield insight as to the system properties that must be considered for future application of the method. In addition, upon establishment of a computationally practicable implementation of the MOF-synthesis simulation method, the method could be applied to different, known, systems of MOFs. This would allow analysis of the capacity of the method in relation to the characteristics of the building blocks, such as the linker flexibility and the metal-source lability. This would yield a thorough evaluation of the MOF-synthesis simulation method, and would give insight as to the level of confidence there could be in the application of the method as a predictive tool.

Appendix A

This section provides the definitions of all acronyms used repeatedly in this thesis.

AIM	Atoms In Molecules
DFT	Density Functional Theory
IAST	Ideal Adsorbed Solution Theory
IRMOF	Isoreticular Metal-Organic Framework
GCMC	Grand Canonical Monte Carlo
HIAST	Heterogeneous Ideal Adsorbed Solution Theory
kMC	kinetic Monte Carlo
LJ	Lennard Jones
MC	Monte Carlo
MD	Molecular Dynamics
MOF	Metal-Organic Framework
MUSIC	Multipurpose Simulation Code
OPLS-AA	Optimised Potentials for Liquid Simulations – All Atom
OPLS-UA	Optimised Potentials for Liquid Simulations – United Atom
PMF	Potential of Mean Force
PSA	Pressure Swing Adsorption
UFF	Universal Force Field
WHAM	Weighted Histogram Analysis Method

Appendix B

This section provides details of the simulation code relevant to the work in this thesis. The source code that was employed in both the development of the implicit-solvent method and the execution of simulations of the synthesis of MOFs is a modified form of the Music simulation code. The Music code is available from http://zeolites.cqe.northwestern.edu/Music/ and the simulation code employed in this work might be obtained by contacting N.Cessford@ed.ac.uk.

B.1 Code alterations and development

Execution of the work in this thesis necessitated the implementation of changes to the Music code in relation to three main purposes – umbrella sampling, the application of PMFs, and MOF-synthesis simulations. The code development associated with these undertakings is described in Sections B.1.1, B.1.2, and B.1.3, respectively. In addition, the code development associated with the WHAM analysis, which was developed to be functional separately from Music, is given in Section B.1.4. As regards evaluation of IAST with respect to MOFs, both single-component and mixture adsorption isotherms were determined using an unmodified version of Music. However, obtaining the associated IAST predictions required code development, as described in Section B.1.5.

B.1.1 Umbrella Sampling

The execution of umbrella sampling simulations required implementation of the capacity to apply a harmonic potential, according to definable parameters, between species. In addition, to enable construction of the probability distributions necessary for analysis of the umbrella sampling simulations, the ability to collect, store and display data relating to the positions of the species required implementation. The introduction of these capabilities necessitated, in addition to alterations to existing modules, the creation of the following new modules in Music:

- umbrellapot.F90
- pcf.F90

B.1.2 PMF Application

The execution of implicit-solvent simulations in which the interactions between solutes were represented by means of PMFs required implementation of the capacity to apply PMFs as potentials between pairs of species, making use of PMF-data supplied in input files. This necessitated, in addition to alterations to existing modules, the creation of the following new module in Music:

- PMF.F90

B.1.3 MOF-Synthesis Simulations

The execution of simulations of the synthesis of MOFs required development of a means of carrying out reaction moves as per the approach outlined in Chapter 5. This necessitated, in addition to alterations to existing modules, the creation of the following new module in Music:

- reaction.F90

B.1.4 WHAM Analysis

Obtaining PMFs from the data obtained from umbrella sampling simulations required development of a code that enables the WHAM analysis in terms of the associated equations, as specified in Chapter 3. This code was not developed as part of Music but requires, as input, the probability distributions produced in umbrella sampling simulations using Music, and produces, as output, the PMF data necessary for implementation of PMFs in Music.

B.1.5 Evaluation of IAST

Evaluating IAST predictions from the relevant single-component isotherm data required development of a code that enabled application of the underlying theory, as specified in Chapter 6. This code was developed separately from Music. As input, this code utilises the single-component isotherms produced in Music simulations and, as output, produces IAST predictions with which mixture simulation data, obtained using Music, were compared.

B.2 Input and Output Files

The simulations carried out in this work are based on the use of the Music simulation code. Details of the input and output files associated with the use of Music can be found at http://zeolites.cqe.northwestern.edu/Music/Documentation/index.html. With the presence of such comprehensive instruction in the use of Music, because the input/output files are extremely extensive, only the alterations that were made to the Music input and output files to enable the execution of the work described herein are given in this thesis.

The execution of the self-assembly simulations described in this thesis required utilisation of the Music files in their standard formats. A brief description of the purposes and contents of these standard input and output files is given in Section B.2.1. The alterations to the input and output files that were necessary for execution

of the umbrella sampling simulations are shown in Section B.2.2 and the file changes necessary for the application of PMFs are given in Section B.2.3. The modifications to the Music input and output files involved in the execution of MOF-synthesis simulations are shown in Section B.2.4.

B.2.1. Self-Assembly Simulations

A number of input files are required for the execution of self-assembly simulations in the NVT ensemble in Music:

- The control file. This file contains details of the user-specified parameters necessary for execution of the required simulation. This file gives information relating to the molecules (and the associated atoms) involved in the simulation, the simulation cell, the moves involved in the simulation, and the initial configuration as well as various other details.
- The molecule-molecule interactions file. This file contains details of the specifications and parameters required for determination of the interactions between the all of the pairs of molecules in the simulation.
- The atom-atom interactions file. This file contains details and the relevant parameters required for representation of the interactions between all of the pairs of atoms in the molecules in the simulation.
- The intramolecular interactions file. This file specifies the intramolecular interactions that must be considered for each molecule involved in the simulation.
- The initial configuration file. This file gives the initial positions of all of the species represented in the simulation.
- Molecule files. These files, of which one is required for each type of molecule involved in the simulation, give details of the structure, charge, connectivity, and intramolecular behaviour of the molecules represented in the simulation.
- Atom files. These files, of which one is required for each type of molecule involved in the simulation, give basic information relating to the atoms in the simulated molecules.

In addition, running a self-assembly simulation in Music produces a number of output files:

- The log file. This file gives details of all of the output from the simulation, including information relating to the acceptance of moves and a summary of the energies associated with the simulation.
- The movie file. This file gives, periodically throughout the simulation, the locations of the atoms in the simulated molecules. This enables visualisation of the behaviour of the molecules in the simulation.

B.2.2. Umbrella Sampling

Development of the Music code to enable the execution of umbrella-sampling simulations required a small number of changes to the input files:

- In the molecule-molecule interactions file, the utilisation of the harmonic potential required to enable umbrella sampling must be specified for the relevant pair of molecules. This can be achieved by adding the keyword UMB to the specification of the non-coulombic (identified by the keyword NCOUL) interaction associated with the pair of molecules between which the constraining umbrella potential is applied. In this case, the umbrella potential is applied between two cobalt ions (identified by the names Cobalt_ion_1 and Cobalt_ion_2).

```
Cobalt_ion_1   Cobalt_ion_2        NCOUL  BASIC LJ  UMB  FAST
```

- In the atom-atom interactions file, the details of the harmonic potential required to enable umbrella sampling must be specified for the relevant pair of atoms. This can be achieved through the addition of a line of keywords equivalent to that given. This specification line relates to the application of an umbrella potential (denoted by UMB) between a pair of cobalt atoms (denoted by Cobalt_1 and Cobalt_2). The DIST keyword and associated parameter (denoted, in this case, by @2.0) relate to the distance at which the harmonic potential is centred, and the FORCE keyword and associated

parameter (denoted, in this case, by @28.0) relate to the force constant associated with the harmonic potential.

```
Cobalt_1     Cobalt_2        UMB DIST@2.0  FORCE@28.0
```

- In the control file, in order to collect the data required to produce the biased probability distributions resulting from the umbrella sampling simulation, the addition of a new section (titled pcf Information) was necessary. This section contains information relating to the atoms between which the umbrella-sampling data must be collected, and also allows definition of the histogram parameters.

```
-------- pcf Information ---------------------------------------------
1                          # histogram number
pcf_25Auc_Co-Co       # output filename
0                  # % data from start to discount
distance           # method of analysis
0, 10, 800         # min, max, no. bins
3, 1              # molecule1 type number, atom number in mol
2, 1              # molecule2 type number, atom number in mol
NULL              # filename to read data from, if required
```

In addition to the standard music output files, umbrella-sampling simulations produce, as output, files displaying the sampled data from which probability distributions were determined. A truncated example is shown below.

```
#Pair Correlation Function results
#------------------------------
#Analysis in terms of: distance
#Number of steps:    100000000
#Distance calculated between: Cobalt_1 & Cobalt_2
   #Bin No.   Distance      Freq.   Norm. Freq.   Cumul.
      1    0.625000E-02        0    0.000000    0.000000
      2    0.187500E-01        0    0.000000    0.000000
      3    0.312500E-01        0    0.000000    0.000000
      4    0.437500E-01        0    0.000000    0.000000
      5    0.562500E-01        0    0.000000    0.000000
      6    0.687500E-01        0    0.000000    0.000000
      7    0.812500E-01        0    0.000000    0.000000
      8    0.937500E-01        0    0.000000    0.000000
      9    0.106250E+00        0    0.000000    0.000000
     10    0.118750E+00        0    0.000000    0.000000
```

B.2.3. PMF Application

Development of the Music code to enable the application of PMFs so as to allow the execution of implicit-solvent simulations required a small number of changes to the input files:

- In the molecule-molecule file, the use of PMFs to represent the interactions between a pair of molecules can be achieved by adding the keyword PMF to the specification of the non-coulombic interaction associated with the pair of molecules.

```
Cobalt_ion_1    Cobalt_ion_2    NCOUL  BASIC  PMF FAST
```

- In the atom-atom interactions file, the details of the applied PMFs must be specified. This can be achieved through the addition a line of keywords equivalent to that shown below. This specification line relates the application of a PMF (denoted by the keyword PMF). The DATA keyword and associated filename relate to the location of the PMF data, the HICUT keyword and associated parameter convey the maximum separation distance at which the PMF potential is applied, and the LOCUT keyword and associated parameter convey the minimum separation distance at which the PMF potential is applied.

```
Cobalt_1  Cobalt_2  PMF  DATA@Co-Co_PMF.dat  HICUT@9.999  LOCUT@2.5875
```

- An input file (named according to the filename given in the atom-atom interactions file) providing the necessary PMF-data is also required. Such files give the PMF values (right column) at distance intervals (left column) along the reaction coordinate.

```
2.58750    91.33183
2.60000    90.59554
2.61250    90.02142
2.62500    89.83972
2.63750    88.61098
2.65000    88.38588
2.66250    87.74408
2.67500    86.92867
   :          :
   :          :
```

B.2.4. MOF-Synthesis Simulation

Development of the Music code to enable the execution of MOF-synthesis simulations required a number of changes to the input files:

- In the control file the addition of a new section was necessary. This section gives information relating to the 'product' molecules in the simulation. The product molecules are all of those that do not exist at the start of the simulation. The final two lines in this section represent the manner in which reaction moves are specified. This manner of specification is also used for the molecule types that are present from the start of the simulation. The final line in this section refers to the name of another input file, in which data relating to the reaction parameters are provided.

```
------ Product Information ------------------------------------------
100.                # Fugacity (kPa)(Range) filename
Null                # sitemap filename (Null = no sitemap)
3                   # no of gcmc movetypes
1.0, 1.0, 1.0       # move type weights
RTRANSLATE          # type of move
0.2, 1              # Delta Translate, adjust delta option (0=NO, 1=YES)
RROTATE             # type of move
0.2, 1              # Delta Translate, adjust delta option (0=NO, 1=YES)
REACT               # type of move
reaction_Prod.dat  # reaction data filename
```

- Parameters relating to the reaction moves in the MOF-synthesis simulations are provided in files (one for each of the starting molecule types, and one for the product molecules) such as that shown below. An example of such a reaction-data file is given below for a product molecule.

```
Number of reactions: 2   # No. reaction types

Reaction #1   FORWARD
OS Co         # Reacting atoms
--            # Lost atoms
2.5           # Reaction Distance, Angstroms
-18.77        # Forward Reaction Enthalpy
2.09          # Forward Reaction Bond Length

Reaction #2   REVERSE
OS Co         # Reacted atoms
18.77         # Reverse Reaction Enthalpy
```

In addition to the standard music output files, MOF-synthesis simulations produce, as output, a file containing details of the coordination of the metal centres represented in the simulation. The coordination data is given periodically, at a specified interval, throughout the simulation. Shown below is a segment of such a file.

Frame1

#Bin No.	Coord. No.	Freq.	Norm. Freq.	Cumul.
1	0	100	1.200000	1.000000
2	1	0	0.000000	1.000000
3	2	0	0.000000	1.000000
4	3	0	0.000000	1.000000
5	4	0	0.000000	1.000000
6	5	0	0.000000	1.000000

Frame2

#Bin No.	Coord. No.	Freq.	Norm. Freq.	Cumul.
1	0	95	1.140000	0.950000
2	1	5	0.060000	1.000000
3	2	0	0.000000	1.000000
4	3	0	0.000000	1.000000
5	4	0	0.000000	1.000000
6	5	0	0.000000	1.000000
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮

B.3 Computer Resources and Simulation Details

All of the simulations involved in the work described in this thesis were carried out using the following resources:

- The CLX compute cluster[230] managed by the School of Engineering at the University of Edinburgh. This cluster is a Beowulf[231] cluster based on 2.4 GHz or 2.6 GHz Advanced Micro Devices (AMD) Opteron dual-core CPUs. Each machine node has 4 CPUs and 32 GB of RAM (8 GB of RAM per CPU). Sun Grid Engine is utilised for managing jobs on this cluster.
- The Condor compute pool[232] in the School of Engineering at the University of Edinburgh. Condor is a batch job management system that schedules jobs submitted to a pool of interconnected computers, taking advantage of the idle CPU-cycles. The Engineering Condor Pool consists of two types of node: a large number of servers with 2 cores and 2 GB of memory (1 GB per core),

and a small number of servers with many cores and a large amount of memory (up to 32 GB).

- The compute component of the Edinburgh Compute and Data Facility (ECDF),[233] known as Eddie. Eddie is composed of two parts: 130 IBM dx360M2 iDataPlex servers with two Intel Westmere E5620 quad-core processors and 24 GB of RAM, connected with gigabit ethernet (Mark 2 Phase 1), and 156 IBM dx360M3 iDataPlex servers with two Intel Westmere E5645 six-core processors and 24 GB of RAM, connected with gigabit ethernet (Mark 2 Phase 2). By default Eddie offers 2 GB to each job slot used, though more is available on request.

The CLX compute cluster was employed to carry out explicit-solvent self-assembly simulations and all simulations involved in the evaluation of IAST with respect to MOFs. As described in Chapter 2, explicit-solvent self-assembly simulations had execution times of the order of weeks. The GCMC simulations carried out as part of the evaluation of IAST had execution times ranging from a few minutes to a number of days, and the simulations performed to evaluate IAST predictions were completed in minutes. The umbrella sampling simulations and WHAM analysis associated with the development of the implicit-solvent method, and all sampling simulations involved in the evaluation of the implicit-solvent method were performed using the Condor Pool. The use of the Condor Pool allowed the simultaneous execution of umbrella sampling simulations in numerous windows, with, for the precision attained in this work, individual simulations lasting in the region of one day. The WHAM analysis of the probability distributions produced in the umbrella sampling simulations required simulation times of the order of minutes. As described in Chapter 4, whilst the implicit-solvent sampling simulations associated with the evaluation of the PMF-approach required only minutes-long simulations, the equivalent explicit-solvent sampling simulations lasted weeks. The Condor Pool was also utilised to carry out the MOF-synthesis simulations, which, as described in Chapter 5, had execution times of a matter of hours. Eddie was employed in the execution of the AIM-analysis simulations, which had simulation times of under an

hour. All of the simulations executed as part of this work were carried out via serial computation.

References

1. The International Energy Agency *Key World Energy Statistics*; France, 2012.

2. Lu, G. Q.; da Costa, J. C. D.; Duke, M.; Giessler, S.; Socolow, R.; Williams, R. H.; Kreutz, T., Inorganic membranes for hydrogen production and purification: A critical review and perspective. *Journal of Colloid and Interface Science* 2007, 314, 589-603.

3. Maxwell, G. R., *Synthetic Nitrogen Products: A Practical Guide to the Products and Processes*. Kluwer Academic/Plenum Publishers: New York, 2004.

4. Ball, M.; Weitschel, M., *The Hydrogen Economy: Opportunities and Challenges*. Cambridge University Press: Cambridge, 2009.

5. Adhikari, S.; Fernando, S., Hydrogen membrane separation techniques. *Industrial & Engineering Chemistry Research* 2006, 45, 875-881.

6. Yampolski, Y.; Freeman, B., *Membrane Gas Separations*. Wiley: Chichester, 2010.

7. Stöcker, J.; Whysall, M.; Miller, G. Q. *30 Years of PSA Technology for Hydrogen Purification*; 1998.

8. Masel, R. I., *Principles of Adsorption and Reaction on Solid Surfaces*. Wiley: New York, 1996.

9. Crittenden, B.; Thomas, W. J., *Adsorption Technology and Design*. Butterworth-Heinemann: Oxford, UK, 1998.

10. Rouquerol, J.; Avnir, D.; Fairbridge, C. W.; Everett, D. H.; Haynes, J. H.; Pernicone, N.; Ramsay, J. D. F.; Sing, K. S. W.; Unger, K. K., Recommendations for the characterisation of porous solids. *Pure and Applied Chemistry* 1994, 66, 1739-1758.

11. Czaja, A. U.; Trukhan, N.; Müller, U., Industrial applications of metal-organic frameworks. *Chemical Society Reviews* 2009, 38, 1284-1293.

12. Kuppler, R. J.; Timmons, D. J.; Fang, Q. R.; Li, J. R.; Makal, T. A.; Young, M. D.; Yuan, D. Q.; Zhao, D.; Zhuang, W. J.; Zhou, H. C., Potential applications of metal-organic frameworks. *Coordination Chemistry Reviews* 2009, 253, 3042-3066.

13. Ferey, G., Hybrid porous solids: past, present, future. *Chemical Society Reviews* 2008, 37, 191-214.

14. James, S. L., Metal-organic frameworks. *Chemical Society Reviews* 2003, 32, 276-288.

15. Rowsell, J. L. C.; Yaghi, O. M., Metal-organic frameworks: a new class of porous materials. *Microporous and Mesoporous Materials* 2004, 73, 3-14.

16. Wright, P. A., *Microporous Framework Solids*. RSC Publishing: Cambridge, UK, 2008.

17. Farha, O. K.; Eryazici, I.; Jeong, N. C.; Hauser, B. G.; Wilmer, C. E.; Sarjeant, A. A.; Snurr, R. Q.; Nguyen, S. T.; Yazaydin, A. O.; Hupp, J. T., Metal-Organic Framework Materials with Ultrahigh Surface Areas: Is the Sky the Limit? *Journal of the American Chemical Society* 2012, 134, 15016-15021.

18. Valtchev, V.; Mintova, S.; Tsapatsis, M., *Ordered Porous Solids: Recent Advances and Prospects*. Elsevier: Oxford, UK, 2009.

19. Maji, T. K.; Kitagawa, S., Chemistry of porous coordination polymers. *Pure and Applied Chemistry* 2007, 79, 2155-2177.

20. Uemura, K.; Matsuda, R.; Kitagawa, S., Flexible microporous coordination polymers. *Journal of Solid State Chemistry* 2005, 178, 2420-2429.

21. Eddaoudi, M.; Kim, J.; Rosi, N.; Vodak, D.; Wachter, J.; O'Keeffe, M.; Yaghi, O. M., Systematic design of pore size and functionality in isoreticular MOFs and their application in methane storage. *Science* 2002, 295, 469-472.

22. Rosi, N. L.; Eckert, J.; Eddaoudi, M.; Vodak, D. T.; Kim, J.; O'Keeffe, M.; Yaghi, O. M., Hydrogen storage in microporous metal-organic frameworks. *Science* 2003, 300, 1127-1129.

23. Rosi, N. L.; Eddaoudi, M.; Kim, J.; O'Keeffe, M.; Yaghi, O. M., Advances in the chemistry of metal-organic frameworks. *Crystengcomm* 2002, 401-404.

24. Kurmoo, M., Magnetic metal-organic frameworks. *Chemical Society Reviews* 2009, 38, 1353-1379.

25. Allendorf, M. D.; Bauer, C. A.; Bhakta, R. K.; Houk, R. J. T., Luminescent metal-organic frameworks. *Chemical Society Reviews* 2009, 38, 1330-1352.

26. MacGillivray, L. R., *Metal-Organic Frameworks: Design and Application*. Wiley: New Jersey, 2010.

27. Rosseinsky, M. J., Recent developments in metal-organic framework chemistry: design, discovery, permanent porosity and flexibility. *Microporous and Mesoporous Materials* 2004, 73, 15-30.

28. Horcajada, P.; Gref, R.; Baati, T.; Allan, P. K.; Maurin, G.; Couvreur, P.; Ferey, G.; Morris, R. E.; Serre, C., Metal-Organic Frameworks in Biomedicine. *Chemical Reviews* 2012, 112, 1232-1268.

29. Stock, N.; Biswas, S., Synthesis of Metal-Organic Frameworks (MOFs): Routes to Various MOF Topologies, Morphologies, and Composites. *Chemical Reviews* 2012, 112, 933-969.

30. Kickelbick, G., *Hybrid Materials: Synthesis, Characterisation and Applications*. Wiley-VCH: Weinheim, 2007.

31. Wang, Z. Q.; Cohen, S. M., Postsynthetic modification of metal-organic frameworks. *Chemical Society Reviews* 2009, 38, 1315-1329.

32. Oganov, A. R., *Modern Methods of Crystal Structure Prediction*. Wiley-VCH: Weinheim, 2011.

33. Price, S. L., From crystal structure prediction to polymorph prediction: interpreting the crystal energy landscape. *Physical Chemistry Chemical Physics* 2008, 10, 1996-2009.

34. Day, G. M.; Cooper, T. G.; Cruz-Cabeza, A. J.; Hejczyk, K. E.; Ammon, H. L.; Boerrigter, S. X. M.; Tan, J. S.; Della Valle, R. G.; Venuti, E.; Jose, J.; Gadre, S. R.; Desiraju, G. R.; Thakur, T. S.; van Eijck, B. P.; Facelli, J. C.; Bazterra, V. E.; Ferraro, M. B.; Hofmann, D. W. M.; Neumann, M. A.; Leusen, F. J. J.; Kendrick, J.; Price, S. L.; Misquitta, A. J.; Karamertzanis, P. G.; Welch, G. W. A.; Scheraga, H. A.; Arnautova, Y. A.; Schmidt, M. U.; van de Streek, J.; Wolf, A. K.; Schweizer, B., Significant progress in predicting the crystal structures of small organic molecules - a report on the fourth blind test. *Acta Crystallographica Section B-Structural Science* 2009, 65, 107-125.

35. Woodley, S. M.; Catlow, R., Crystal structure prediction from first principles. *Nature Materials* 2008, 7, 937-946.

36. Kirkpatrick, S.; Gelatt, C. D.; Vecchi, M. P., OPTIMIZATION BY SIMULATED ANNEALING. *Science* 1983, 220, 671-680.

37. Wales, D. J.; Scheraga, H. A., Review: Chemistry - Global optimization of clusters, crystals, and biomolecules. *Science* 1999, 285, 1368-1372.

38. Leszczynski, J., *Handbook of Computational Chemistry*. Springer: London, 2012.

39. Woodley, S. M., Prediction of crystal structures using evolutionary algorithms and related techniques. *Applications of Evolutionary Computation in Chemistry* 2004, 110, 95-132.

40. Johnston, R. L., Evolving better nanoparticles: Genetic algorithms for optimising cluster geometries. *Dalton Transactions* 2003, 4193-4207.

41. Mingos, D. M. P., *Applications of Evolutionary Computation in Chemistry*. Springer: 2004.

42. Lommerse, J. P. M.; Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Gavezzotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Mooij, W. T. M.; Price, S. L.; Schweizer, B.; Schmidt, M. U.; van Eijck, B. P.; Verwer, P.; Williams, D. E., A test of crystal structure prediction of small organic molecules. *Acta Crystallographica Section B-Structural Science* 2000, 56, 697-714.

43. Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Dzyabchenko, A.; Erk, P.; Gavezzotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Lommerse, J. P. M.; Mooij, W. T. M.; Price, S. L.; Scheraga, H.; Schweizer, B.; Schmidt, M. U.; van Eijck, B. P.; Verwer, P.; Williams, D. E., Crystal structure prediction of small organic molecules: a second blind test. *Acta Crystallographica Section B-Structural Science* 2002, 58, 647-661.

44. Day, G. M.; Motherwell, W. D. S.; Ammon, H. L.; Boerrigter, S. X. M.; Della Valle, R. G.; Venuti, E.; Dzyabchenko, A.; Dunitz, J. D.; Schweizer, B.; van Eijck, B. P.; Erk, P.; Facelli, J. C.; Bazterra, V. E.; Ferraro, M. B.; Hofmann, D. W. M.; Leusen, F. J. J.; Liang, C.; Pantelides, C. C.; Karamertzanis, P. G.; Price, S. L.; Lewis, T. C.; Nowell, H.; Torrisi, A.; Scheraga, H. A.; Arnautova, Y. A.; Schmidt, M. U.; Verwer, P., A third blind test of crystal structure prediction. *Acta Crystallographica Section B-Structural Science* 2005, 61, 511-527.

45. Bardwell, D. A.; Adjiman, C. S.; Arnautova, Y. A.; Bartashevich, E.; Boerrigter, S. X. M.; Braun, D. E.; Cruz-Cabeza, A. J.; Day, G. M.; Della Valle, R. G.; Desiraju, G. R.; van Eijck, B. P.; Facelli, J. C.; Ferraro, M. B.; Grillo, D.; Habgood, M.; Hofmann, D. W. M.; Hofmann, F.; Jose, K. V. J.; Karamertzanis, P. G.; Kazantsev, A. V.; Kendrick, J.; Kuleshova, L. N.; Leusen, F. J. J.; Maleev, A. V.; Misquitta, A. J.; Mohamed, S.; Needs, R. J.; Neumann, M. A.; Nikylov, D.; Orendt, A. M.; Pal, R.; Pantelides, C. C.; Pickard, C. J.; Price, L. S.; Price, S. L.; Scheraga, H. A.; van de Streek, J.; Thakur, T. S.; Tiwari, S.; Venuti, E.; Zhitkov, I. K., Towards crystal structure prediction of complex organic compounds - a report on the fifth blind test. *Acta Crystallographica Section B-Structural Science* 2011, 67, 535-551.

46. Draznieks, C. M.; Newsam, J. M.; Gorman, A. M.; Freeman, C. M.; Ferey, G., De novo prediction of inorganic structures developed through automated assembly of secondary building units (AASBU method). *Angewandte Chemie-International Edition* 2000, 39, 2270-2275.

47. Kim, E.; Whitesides, G. M., Use of Minimal Free-Energy and Self-Assembly to Form Shapes. *Chemistry of Materials* 1995, 7, 1257-1264.

48. Whitesides, G. M.; Boncheva, M., Beyond molecules: Self-assembly of mesoscopic and macroscopic components. *P Natl Acad Sci USA* 2002, 99, 4769-4774.

49. Olenyuk, B.; Whiteford, J. A.; Fechtenkotter, A.; Stang, P. J., Self-assembly of nanoscale cuboctahedra by coordination chemistry. *Nature* 1999, 398, 796-799.

50. van Gunsteren, W. F.; Berendsen, H. J. C., Computer-Simulation of Molecular-Dynamics - Methodology, Applications, and Perspectives in Chemistry. *Angew Chem Int Edit* 1990, 29, 992-1023.

51. McQuarrie, D. A., *Statistical Mechanics*. Harper & Row: New York, 1976.

52. Chandler, D., *Introduction to Modern Statistical Mechanics*. Oxford University Press: New York, 1987.

53. Hill, T. L., *Introduction to statistical thermodynamics*. Dover Publications: New York, 1986.

54. Prohofsky, E., *Statistical Mechanics and Stability of Macromolecules, Application to Bond Disruption, Base-pair Separation, Melting, and Drug Dissociation of the DNA Double Helix.* Cambridge University Press: New York, 1995.

55. Pasini, P.; Zannoni, C.; Zumer, S., *Computer Simulations of Liquid Crystals and Polymers*. Kluwer Academic Publishers: Amsterdam, 2005.

56. Frenkel, D.; Smit, B., *Understanding Molecular Simulation, From Algorithms to Applications*. 2nd ed.; Academic Press: 2002.

57. Allen, M. P.; Tildesley, D. J., *Computer Simulation of Liquids*. Clarendon Press: Oxford, 1987.

58. Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E., Equation of state calculations by fast computing machines. *Journal of Chemical Physics* 1953, 21, 1087-1092.

59. Tuckerman, M. E., *Statistical Mechanics: Theory and Molecular Simulation*. Oxford university Press: Oxford, 2010.

60. Rimai, D. S.; DeMejo, L. P.; Mittal, K. L., *Fundamentals of Adhesion and Interfaces*. VSP: Utrecht, 1995.

61. Becker, O. M.; MacKerell, A. D.; Roux, B.; Watanabe, M., *Computational Biochemistry and Biophysics*. Marcel Dekker, Inc.: New York, 2001.

62. Lucas, K., *Molecular Models for Fluids*. Cambridge University Press: London, 2007.

63. da Silva, L. F. M.; Ochsner, A.; Adams, R. D., *Handbook of Adhesion Technology*. Springer: 2011; Vol. 1.

64. Letcher, T., *Chemical Thermodynamics for Industry*. The Royal Society of Chemistry: Cambridge, 2004.

65. Ungerer, P.; Tavitian, B.; Boutin, A., Applications of Molecular Simulation in the Oil and Gas Industry, Monte Carlo Methods. Editions Technip: Paris, 2005.

66. Schlick, T., *Molecular Modelling and Simulation: An Interdisciplinary Guide*. Springer: New York, 2010.

67. Bultinck, P.; De Winter, H.; Langenaeker, W.; Tollenaere, J. P., *Computational Medicinal Chemistry for Drug Discovery*. Marcell Dekker, Inc.: New York, 2004.

68. Feller, S. E., *Computational Modeling of Membrane Bilayers, Current Topics in Membranes*. Academic Press: London, 2008; Vol. 60.

69. Ewald, P. P., The calculation of optical and electrostatic grid potential. *Ann. Phys.-Berlin* 1921, 64, 253-287.

70. Fehske, H.; Schneider, R.; Weiβe, A., *Computational Many-Particle Physics, Lecture Notes in Physics*. Springer: Berlin, 2008; Vol. 739.

71. Schwede, T.; Peitsch, M., *Computational Structural Biology, Methods and Applications*. World Scientific: Singapore, 2008.

72. Sagui, C.; Darden, T. A., Molecular dynamics simulations of biomolecules: Long-range electrostatic effects. *Annual Review of Biophysics and Biomolecular Structure* 1999, 28, 155-179.

73. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A smooth particle mesh Ewald method. *Journal of Chemical Physics* 1995, 103, 8577-8593.

74. Wolf, D.; Keblinski, P.; Phillpot, S. R.; Eggebrecht, J., Exact method for the simulation of Coulombic systems by spherically truncated, pairwise r(-1) summation. *Journal of Chemical Physics* 1999, 110, 8254-8282.

75. Adams, D. J., On the use of the Ewald summation in computer-simulation. *Journal of Chemical Physics* 1983, 78, 2585-2590.

76. Gdoutos, E. E.; Agrawal, R.; Espinosa, H. D., Comparison of the Ewald and Wolf methods for modeling electrostatic interactions in nanowires. *International Journal for Numerical Methods in Engineering* 2010, 84, 1541-1551.

77. Ma, Y.; Garofalini, S. H., Modified Wolf electrostatic summation: Incorporating an empirical charge overlap. *Molecular Simulation* 2005, 31, 739-748.

78. Forster, P. M.; Burbank, A. R.; Livage, C.; Ferey, G.; Cheetham, A. K., The role of temperature in the synthesis of hybrid inorganic-organic materials: the example of cobalt succinates. *Chemical Communications* 2004, 368-369.

79. Forster, P. M.; Stock, N.; Cheetham, A. K., A high-throughput investigation of the role of pH, temperature, concentration, and time on the synthesis of hybrid inorganic-organic materials. *Angewandte Chemie-International Edition* 2005, 44, 7608-7611.

80. Livage, C.; Egger, C.; Ferey, G., Hydrothermal versus nonhydrothermal synthesis for the preparation of organic-inorganic solids: The example of cobalt(II) succinate. *Chemistry of Materials* 2001, 13, 410-414.

81. Livage, C.; Egger, C.; Ferey, G., Hybrid open networks (MIL 16): Synthesis, crystal structure, and ferrimagnetism of Co-4(OH)(2)(H2O)(2)(C4H4O4)(3)·2H(2)O, a new layered cobalt(II) carboxylate with 14-membered ring channels. *Chemistry of Materials* 1999, 11, 1546-1550.

82. Long, L. S.; Chen, X. M.; Tong, M. L.; Sun, Z. G.; Ren, Y. P.; Huang, R. B.; Zheng, L. S., A unique open inorganic-organic framework with alternate hexa- and penta-coordinate cobalt(II) sites. Synthesis, crystal structure and magnetic properties of [Co3(C4H4O4)2.5(OH)]n·0.5nH2O. *Journal of the Chemical Society-Dalton Transactions* 2001, 2888-2890.

83. Livage, C.; Egger, C.; Nogues, M.; Ferey, G., Hybrid open frameworks (MIL-n). Part 5 - Synthesis and crystal structure of MIL-9: a new three-dimensional ferrimagnetic cobalt(II) carboxylate with a two-dimensional array of edge-sharing Co octahedra with 12-membered rings. *Journal of Materials Chemistry* 1998, 8, 2743-2747.

84. Forster, P. M.; Burbank, A. R.; O' Sullivan, M. C.; Guillou, N.; Livage, C.; Ferey, G.; Stock, N.; Cheetham, A. K., Single-crystal characterization of Co-7(OH)(6)(H2O)(3)(C4H4O4)(4)·7H(2)O; A new cobalt succinate identified through high-throughput synthesis. *Solid State Sciences* 2005, 7, 1549-1555.

85. Forster, P. M.; Tafoya, M. M.; Cheetham, A. K., Synthesis and characterization of Co-7(OH)(12)(C2H4S2O6)(H2O)(2) - a single crystal structural study of a ferrimagnetic layered cobalt hydroxide. *Journal of Physics and Chemistry of Solids* 2004, 65, 11-16.

86. Mahata, P.; Prabu, M.; Natarajan, S., Role of temperature and time in the formation of infinite -M-O-M-linkages and isolated clusters in MOFs: A few illustrative examples. *Inorganic Chemistry* 2008, 47, 8451-8463.

87. Jorgensen, W. L.; Gao, J., Monte-Carlo Simulations of the Hydration of Ammonium and Carboxylate Ions. *Journal of Physical Chemistry* 1986, 90, 2174-2182.

88. Price, D. J.; Roberts, J. D.; Jorgensen, W. L., Conformational complexity of succinic acid and its monoanion in the gas phase and in solution: Ab initio calculations and Monte Carlo simulations. *Journal of the American Chemical Society* 1998, 120, 9672-9679.

89. Jorgensen, W. L.; Madura, J. D.; Swenson, C. J., Optimized Intermolecular Potential Functions for Liquid Hydrocarbons. *Journal of the American Chemical Society* 1984, 106, 6638-6646.

90. Jorgensen, W. L.; Swenson, C. J., Optimized Intermolecular Potential Functions for Amides and Peptides - Structure and Properties of Liquid Amides. *Journal of the American Chemical Society* 1985, 107, 569-578.

91. Jorgensen, W. L.; Swenson, C. J., Optimized Intermolecular Potential Functions for Amides and Peptides - Hydration of Amides. *Journal of the American Chemical Society* 1985, 107, 1489-1496.

92. Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J., Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society* 1996, 118, 11225-11236.

93. Babu, C. S.; Lim, C., Empirical force fields for biologically active divalent metal cations in water. *J Phys Chem A* 2006, 110, 691-699.

94. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of Simple Potential Functions for Simulating Liquid Water. *Journal of Chemical Physics* 1983, 79, 926-935.

95. Jorgensen, W. L.; Tirado-Rives, J., The OPLS Potential Functions for Proteins - Energy Minimizations for Crystals of Cyclic-Peptides and Crambin. *Journal of the American Chemical Society* 1988, 110, 1657-1666.

96. Gupta, A.; Chempath, S.; Sanborn, M. J.; Clark, L. A.; Snurr, R. Q., Object-oriented programming paradigms for molecular modeling. *Mol. Simul.* 2003, 29, 29-46.

97. Smith, P. E.; Pettitt, B. M., Modeling Solvent in Biomolecular Systems. *J Phys Chem-Us* 1994, 98, 9700-9711.

98. Roux, B.; Simonson, T., Implicit solvent models. *Biophysical Chemistry* 1999, 78, 1-20.

99. Feig, M.; Brooks, C. L., Recent advances in the development and application of implicit solvent models in biomolecule simulations. *Current Opinion in Structural Biology* 2004, 14, 217-224.

100. Orozco, M.; Luque, F. J., Theoretical methods for the description of the solvent effect in biomolecular systems. *Chemical Reviews* 2000, 100, 4187-4225.

101. Eisenberg, D.; Mclachlan, A. D., Solvation Energy in Protein Folding and Binding. *Nature* 1986, 319, 199-203.

102. Fraternali, F.; van Gunsteren, W. F., An efficient mean solvation force model for use in molecular dynamics simulations of proteins in aqueous solution. *Journal of Molecular Biology* 1996, 256, 939-948.

103. Baker, N. A.; Wagoner, J. A., Assessing implicit models for nonpolar mean solvation forces: The importance of dispersion and volume terms. *P Natl Acad Sci USA* 2006, 103, 8331-8336.

104. Kang, Y. K.; Nemethy, G.; Scheraga, H. A., Free-Energies of Hydration of Solute Molecules .1. Improvement of the Hydration Shell-Model by Exact Computations of Overlapping Volumes. *Journal of Physical Chemistry* 1987, 91, 4105-4109.

105. Hopfinger, A. J., Polymer-Solvent Interactions for Homopolypeptides in Aqueous Solution. *Macromolecules* 1971, 4, 731-737.

106. Pratt, L. R.; Chandler, D., Theory of Hydrophobic Effect. *Journal of Chemical Physics* 1977, 67, 3683-3704.

107. Pettitt, B. M.; Rossky, P. J., Alkali-Halides in Water - Ion-Solvent Correlations and Ion-Ion Potentials of Mean Force at Infinite Dilution. *Journal of Chemical Physics* 1986, 84, 5836-5844.

108. Pratt, L. R.; Pohorille, A., Hydrophobic effects and modeling of biophysical aqueous solution interfaces. *Chem Rev* 2002, 102, 2671-2691.

109. Pettitt, B. M.; Howard, J. J.; Perkyns, J. S.; Choudhury, N., Integral Equation Study of the Hydrophobic Interaction between Graphene Plates. *J. Chem. Theory Comput.* 2008, 4, 1928-1939.

110. Kirkwood, J. G., Statistical mechanics of fluid mixtures. *Journal of Chemical Physics* 1935, 3, 300-313.

111. Zwanzig, R. W., High-Temperature Equation of State by a Perturbation Method .1. Nonpolar Gases. *Journal of Chemical Physics* 1954, 22, 1420-1426.

112. Kollman, P., Free-Energy Calculations - Applications to Chemical and Biochemical Phenomena. *Chemical Reviews* 1993, 93, 2395-2417.

113. Mezei, M.; Beveridge, D. L., Free energy simulations. *Annals of the New York Academy of Sciences* 1986, 482, 1-23.

114. Beveridge, D. L.; DiCapua, F. M., Free-Energy Via Molecular Simulation - Applications to Chemical and Biomolecular Systems. *Annual Review of Biophysics and Biophysical Chemistry* 1989, 18, 431-492.

115. Mruzik, M. R.; Abraham, F. F.; Schreiber, D. E.; Pound, G. M., A Monte-Carlo Study of Ion-Water Clusters. *Journal of Chemical Physics* 1976, 64, 481-491.

116. Torrie, G. M.; Valleau, J. P., Non-Physical Sampling Distributions in Monte-Carlo Free-Energy Estimation - Umbrella Sampling. *J Comput Phys* 1977, 23, 187-199.

117. Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M., The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules .1. The Method. *J Comput Chem* 1992, 13, 1011-1021.

118. Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A., Multidimensional Free-Energy Calculations Using the Weighted Histogram Analysis Method. *J Comput Chem* 1995, 16, 1339-1350.

119. Canuto, S., *Solvation Effects on Molecules and Biomolecules: Computational Methods and Applications*. 1st ed.; Springer: 2008.

120. Johnson, M.; Holt, J. M.; Ackers, G. K., *Biothermodynamics, Part B: Methods in Enzymology*. 1st ed.; Academic Press: San Diego, 2009.

121. Lelievre, T.; Rousset, M.; Stoltz, G., *Free Energy Computations: A Mathematical Perspective*. 1st ed.; Imperial College Press: London, 2010.

122. Czaplewski, C.; Kalinowski, S.; Liwo, A.; Scheraga, H. A., Comparison of two approaches to potential of mean force calculations of hydrophobic association: particle insertion and weighted histogram analysis methods. *Molecular Physics* 2005, 103, 3153-3167.

123. Torrie, G. M.; Valleau, J. P., Monte Carlo free energy estimates using non-Boltzmann sampling: application to the sub-critical Lennard-Jones fluid. *Chemical Physics Letters* 1974, 28, 578-581.

124. Valleau, J. P.; Torrie, G. M., A guide to Monte Carlo for statistical mechanics: 2. Byways. In *Statistical Mechanics*, Berne, B. J., Ed. Springer US: New York, 1977; pp 169-194.

125. Rajamani, R.; Naidoo, K. J.; Gao, J., Implementation of an adaptive umbrella sampling method for the calculation of multidimensional potential of mean force of chemical reactions in solution. *Journal of Computational Chemistry* 2003, 24, 1775-1781.

126. Shell, M. S.; Panagiotopoulos, A.; Pohorille, A., Methods based on probability distributions and histograms. In *Free Energy Calculations: Theory and Applications in Chemistry and Biology, of Springer Series in Chemical Physics*, Chipot, C.; Pohorille, A., Eds. Springer: Berlin, 2007; Vol. 86, pp 77-118.

127. Villa, A.; Peter, C.; van der Vegt, N. F. A., Self-assembling dipeptides: conformational sampling in solvent-free coarse-grained simulation. *Phys Chem Chem Phys* 2009, 11, 2077-2086.

128. Shinto, H.; Morisada, S.; Miyahara, M.; Higashitani, K., Langevin dynamics simulations of cationic surfactants in aqueous solutions using potentials of mean force. *Langmuir* 2004, 20, 2017-2025.

129. Zahn, D.; Schilling, B.; Kast, S. M., Enhancement of the wolf damped Coulomb potential: Static, dynamic, and dielectric properties of liquid water from molecular simulation. *Journal of Physical Chemistry B* 2002, 106, 10725-10732.

130. Bogusz, S.; Cheatham, T. E.; Brooks, B. R., Removal of pressure and free energy artifacts in charged periodic systems via net charge corrections to the Ewald potential. *Journal of Chemical Physics* 1998, 108, 7070-7084.

131. Friedman, R. A.; Mezei, M., The Potentials of Mean Force of Sodium-Chloride and Sodium Dimethylphosphate in Water - an Application of Adaptive Umbrella Sampling. *Journal of Chemical Physics* 1995, 102, 419-426.

132. Belch, A. C.; Berkowitz, M.; Mccammon, J. A., Solvation Structure of a Sodium-Chloride Ion-Pair in Water. *Journal of the American Chemical Society* 1986, 108, 1755-1761.

133. Dang, L. X.; Pettitt, B. M.; Rossky, P. J., On the Correlation between Like Ion-Pairs in Water. *Journal of Chemical Physics* 1992, 96, 4046-4047.

134. Guardia, E.; Padro, J. A., On the influence of the ionic charge on the mean force potential of ion-pairs in water. *Journal of Chemical Physics* 1996, 104, 7219-7222.

135. Gervasio, F. L.; Chelli, R.; Marchi, M.; Procacci, P.; Schettino, V., Determination of the potential of mean force of aromatic amino acid complexes in various solvents using molecular dynamics simulations: The case of the tryptophan-histidine pair. *Journal of Physical Chemistry B* 2001, 105, 7835-7846.

136. Rozanska, X.; Chipot, C., Modeling ion-ion interaction in proteins: A molecular dynamics free energy calculation of the guanidinium-acetate association. *Journal of Chemical Physics* 2000, 112, 9691-9694.

137. Bergdorf, M.; Peter, C.; Hunenberger, P. H., Influence of cut-off truncation and artificial periodicity of electrostatic interactions in molecular simulations of solvated ions: A continuum electrostatics study. *Journal of Chemical Physics* 2003, 119, 9129-9144.

138. Demontis, P.; Spanu, S.; Suffritti, G. B., Application of the Wolf method for the evaluation of Coulombic interactions to complex condensed matter systems: Aluminosilicates and water. *Journal of Chemical Physics* 2001, 114, 7980-7988.

139. Rustad, J. R.; Rosso, K. M.; Felmy, A. R., Molecular dynamics investigation of ferrous-ferric electron transfer in a hydrolyzing aqueous solution: Calculation of the pH dependence of the diabatic transfer barrier and the potential of mean force. *Journal of Chemical Physics* 2004, 120, 7607-7615.

140. Dang, L. X.; Pettitt, B. M., A Theoretical-Study of Like Ion-Pairs in Solution. *J Phys Chem-Us* 1990, 94, 4303-4308.

141. Dang, L. X.; Pettitt, B. M., Chloride-Ion Pairs in Water. *Journal of the American Chemical Society* 1987, 109, 5531-5532.

142. Fennell, C. J.; Bizjak, A.; Vlachy, V.; Dill, K. A., Ion Pairing in Molecular Simulations of Aqueous Alkali Halide Solutions. *Journal of Physical Chemistry B* 2009, 113, 6782-6791.

143. Wang, H.; Stalnaker, J.; Chevreau, H.; Lewis, J. P., Potential of mean force calculations using ab initio tight-binding molecular dynamics: Application to N-NO(2) bond dissociation in DMNA and HMX. *Chemical Physics Letters* 2008, 457, 26-30.

144. Fennell, C. J.; Gezelter, J. D., Is the Ewald summation still necessary? Pairwise alternatives to the accepted standard for long-range electrostatics. *Journal of Chemical Physics* 2006, 124.

145. Allen, T. W.; Andersen, O. S.; Roux, B., Ion permeation through a narrow channel: Using gramicidin to ascertain all-atom molecular dynamics potential of mean force methodology and biomolecular force fields. *Biophys J* 2006, 90, 3447-3468.

146. Madhusoodanan, M.; Tembe, B. L., Potential of Mean Force of the Sodium-Chloride Ion-Pair in Dimethyl-Sulfoxide - a Constrained Molecular-Dynamics Study. *J Phys Chem-Us* 1994, 98, 7090-7094.

147. Efron, B., 1977 Rietz Lecture - Bootstrap methods - another look at the jackknife. *Annals of Statistics* 1979, 7, 1-26.

148. Hub, J. S.; de Groot, B. L.; van der Spoel, D., g_wham-A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates. *J. Chem. Theory Comput.* 2010, 6, 3713-3720.

149. Israelachvili, J. N., *Intermolecular and Surface Forces*. 2nd ed.; Academic Press: San Diego, 1991.

150. Santiso, E. E.; Gubbins, K. E., Multi-scale molecular modeling of chemical reactivity. *Molecular Simulation* 2004, 30, 699-748.

151. Ferrario, M., et al., *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology*. Springer: Berlin, 2006; Vol. Volume 1.

152. Hartree, D. R., *The Calculation of Atomic Structure*. Wiley: New York, 1961.

153. Hartree, D. R.; Hartree, W., Self-consistent field, with exchange, for cu(+). *Proceedings of the Royal Society of London Series a-Mathematical and Physical Sciences* 1936, 157, 0490-0502.

154. Moller, C.; Plesset, M. S., Note on an approximation treatment for many-electron systems. *Physical Review* 1934, 46, 0618-0622.

155. Field, M. J., *A Practical Introduction to the Simulation of Molecular Systems*. 2nd Edition ed.; Cambridge University Press: New York, 2007.

156. Deák, P.; Frauenheim, T.; Pederson, M. R., *Computer Simulations of Materials at Atomic Level*. 1st Edition ed.; Wiley-VCH Verlag GmbH: Berlin, 2000.

157. Kohn, W.; Becke, A. D.; Parr, R. G., Density functional theory of electronic structure. *Journal of Physical Chemistry* 1996, 100, 12974-12980.

158. Reimers, J. R., *Computational Approaches for Large Systems: Electronic Structure Approaches for Biotechnology and Nanotechnology*. Wiley: New Jersey, 2011.

159. Stewart, J. J. P., Optimization of parameters for semiempirical methods. 1. Method. *Journal of Computational Chemistry* 1989, 10, 209-220.

160. Stewart, J. J. P., Optimization of parameters for semiempirical methods. 2. Applications. *Journal of Computational Chemistry* 1989, 10, 221-264.

161. Bredow, T.; Jug, K., Theory and range of modern semiempirical molecular orbital methods. *Theoretical Chemistry Accounts* 2005, 113, 1-14.

162. Lewars, E. G., *Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics*. 2nd ed.; Springer: New York, 2011.

163. Ross, R. B.; Mohanty, S., *Multiscale Simulation Methods for Nanomaterials*. Wiley: New Jersey, 2008.

164. Tapanes, N. C. O.; Aranda, D. A. G.; de Mesquita Carneiro, J. W.; Antunes, O. A. C., Transesterification of Jatropha curcas oil glycerides: Theoretical and experimental studies of biodiesel reaction. *Fuel* 2008, 87, 2286-2295.

165. Tsukamoto, K.; Nakamura, S.; Shimizu, K., SAM1 semiempirical calculations on the catalytic cycle of nitric oxide reductase from Fusarium oxysporum. *Journal of Molecular Structure-Theochem* 2003, 624, 309-322.

166. Dykstra, C. E.; Frenking, G.; Kim, K. S.; Scuseria, G. E., *Theory and Applications of Computational Chemistry: The First Forty Years*. Elsevier: Amsterdam, 2005.

167. Clark, T.; Banting, L., *Drug Design Strategies: Computational Techniques and Applications*. The Royal Society of Chemistry: Cambridge, 2012.

168. Bhushan, B., *Springer Handbook of Nanotechnology*. 2nd ed.; Springer: Berlin, 2007.

169. Smith, W. R.; Triska, B., The reaction ensemble method for the computer-simulation of chemical and phase-equilibria. 1. Theory and basic examples. *Journal of Chemical Physics* 1994, 100, 3019-3027.

170. Johnson, J. K.; Panagiotopoulos, A. Z.; Gubbins, K. E., Reactive canonical Monte Carlo - a new simulation technique for reacting or associating fluids. *Molecular Physics* 1994, 81, 717-733.

171. Turner, C. H.; Brennan, J. K.; Lisal, M.; Smith, W. R.; Johnson, J. K.; Gubbins, K. E., Simulation of chemical reaction equilibria by the reaction ensemble Monte Carlo method: a review. *Molecular Simulation* 2008, 34, 119-146.

172. van Duin, A. C. T.; Dasgupta, S.; Lorant, F.; Goddard, W. A., ReaxFF: A reactive force field for hydrocarbons. *Journal of Physical Chemistry A* 2001, 105, 9396-9409.

173. Brenner, D. W.; Shenderova, O. A.; Harrison, J. A.; Stuart, S. J.; Ni, B.; Sinnott, S. B., A second-generation reactive empirical bond order (REBO) potential energy expression for hydrocarbons. *Journal of Physics-Condensed Matter* 2002, 14, 783-802.

174. Chenoweth, K.; van Duin, A. C. T.; Goddard, W. A., ReaxFF reactive force field for molecular dynamics simulations of hydrocarbon oxidation. *Journal of Physical Chemistry A* 2008, 112, 1040-1053.

175. Cranford, S. W.; Buehler, M. J., Biomateriomics. In *Springer Series in Material Science*, Hull, R.; Jagadish, C.; Osgood, R. M.; Parisi, J.; Wang, Z., Eds. Springer: 2012.

176. Fichthorn, K. A.; Weinberg, W. H., Theoretical foundations of dynamic Monte-Carlo simulations. *Journal of Chemical Physics* 1991, 95, 1090-1096.

177. Theodorou, D. N., Progress and Outlook in Monte Carlo Simulations. *Industrial & Engineering Chemistry Research* 2010, 49, 3047-3058.

178. LeSar, R., *Introduction to Computational Materials Science: Fundamentals to Applications*. Cambridge University Press: New York, 2013.

179. Duke, C. B.; Plummer, E. W., *Frontiers in Surface Science and Interface Science*. Elsevier Science: Amsterdam, 2002.

180. Eyring, H., The activated complex in chemical reactions. *Journal of Chemical Physics* 1935, 3, 107-115.

181. Petersson, G. A., Perspective on "The activated complex in chemical reactions" - Eyring H (1935) J Chem Phys 3 : 107. *Theoretical Chemistry Accounts* 2000, 103, 190-195.

182. Warshel, A.; Levitt, M., Theoretical studies of enzymic reactions - dielectric, electrostatic and steric stabilization of carbonium-ion in reaction of lysozome. *Journal of Molecular Biology* 1976, 103, 227-249.

183. Balbuena, P. B.; Seminario, J. M., *Molecular Dynamics: From Classical to Quantum Methods*. Elsevier: Amsterdam, 1999.

184. Amara, P.; Field, M. J., Evaluation of an ab initio quantum mechanical/molecular mechanical hybrid-potential link-atom method. *Theoretical Chemistry Accounts* 2003, 109, 43-52.

185. Amara, P.; Field, M. J., Combined Quantum Mechanical and Molecular Mechanical Potentials. In *Encyclopedia of Computational Chemistry*, John Wiley & Sons, Ltd.: Chichester, 2002.

186. Bader, R. F. W., *Atoms in Molecules: A Quantum Theory*. Oxford University Press: New York, 1994.

187. Rankin, D. W. H.; Mitzel, N. W.; Morrison, C. A., *Structural Methods in Molecular Inorganic Chemistry*. Wiley: Chichester, 2013.

188. Bader, R. F. W., Atoms in molecules. *Accounts of Chemical Research* 1985, 18, 9-15.

189. Popelier, P. L. A., On the full topology of the Laplacian of the electron density. *Coordination Chemistry Reviews* 2000, 197, 169-189.

190. Lewis, G. N., The atom and the molecule. *Journal of the American Chemical Society* 1916, 38, 762-785.

191. Bansal, R. K., *Organic Reaction Mechanisms*. 3rd ed.; Tata McGraw-Hill: New Delhi, 2006.

192. Anslyn, E. V.; Dougherty, D. A., *Modern Physical Organic Chemistry*. University Science Books: California, 2006.

193. Wang, F.; Zhang, Y. H.; Zhao, L. J.; Zhang, H.; Cheng, H.; Shou, J. J., Micro-Raman study on the conformation behavior of succinate in supersaturated sodium succinate aerosols. *Physical Chemistry Chemical Physics* 2008, 10, 4154-4158.

194. Dessent, C. E. H.; Rigby, C., Exploring the microscopic solvation of doubly charged anions: symmetric or asymmetric solvation in the CO2-(CH2)4-CO22-·(H2O)2 dicarboxylate dianion cluster? *Chemical Physics Letters* 2003, 370, 52-61.

195. Kaminski, G.; Duffy, E. M.; Matsui, T.; Jorgensen, W. L., Free energies of hydration and pure liquid properties of hydrocarbons from the OPLS all-atom model. *Journal of Physical Chemistry* 1994, 98, 13077-13082.

196. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J., J. A.,; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision C.01*, Gaussian, Inc.: Wallingford CT, 2009.

197. Mueller, U.; Schubert, M.; Teich, F.; Puetter, H.; Schierle-Arndt, K.; Pastre, J., Metal-organic frameworks - prospective industrial applications. *Journal of Materials Chemistry* 2006, 16, 626-636.

198. Talu, O., Needs, status, techniques and problems with binary gas adsorption experiments. *Adv. Colloid Interface Sci.* 1998, 77, 227-269.

199. Murthi, M.; Snurr, R. Q., Effects of molecular siting and adsorbent heterogeneity on the ideality of adsorption equilibria. *Langmuir* 2004, 20, 2489-2497.

200. Keskin, S.; Liu, J.; Rankin, R. B.; Johnson, J. K.; Sholl, D. S., Progress, Opportunities, and Challenges for Applying Atomically Detailed Modeling to Molecular Adsorption and Transport in Metal-Organic Framework Materials. *Industrial & Engineering Chemistry Research* 2009, 48, 2355-2371.

201. Yang, Q. Y.; Zhong, C. L., Molecular simulation of carbon dioxide/methane/hydrogen mixture adsorption in metal-organic frameworks. *J. Phys. Chem. B* 2006, 110, 17776-17783.

202. Babarao, R.; Hu, Z. Q.; Jiang, J. W.; Chempath, S.; Sandler, S. I., Storage and separation of CO2 and CH4 in silicalite, C-168 schwarzite, and IRMOF-1: A comparative study from monte carlo simulation. *Langmuir* 2007, 23, 659-666.

203. Keskin, S.; Liu, J. C.; Johnson, J. K.; Sholl, D. S., Testing the accuracy of correlations for multicomponent mass transport of adsorbed gases in metal-organic frameworks: Diffusion of H-2/CH4 mixtures in CuBTC. *Langmuir* 2008, 24, 8254-8261.

204. Bae, Y. S.; Mulfort, K. L.; Frost, H.; Ryan, P.; Punnathanam, S.; Broadbelt, L. J.; Hupp, J. T.; Snurr, R. Q., Separation of CO2 from CH4 using mixed-ligand metal-organic frameworks. *Langmuir* 2008, 24, 8592-8598.

205. Liu, B.; Yang, Q.; Xue, C.; Zhong, C.; Chen, B.; Smit, B., Enhanced adsorption selectivity of hydrogen/methane mixtures in metal-organic frameworks with interpenetration: A molecular simulation study. *J. Phys. Chem. C* 2008, 112, 9854-9860.

206. Yang, Q. Y.; Xue, C. Y.; Zhong, C. L.; Chen, J. F., Molecular simulation of separation of CO2 from flue gases in Cu-BTC metal-organic framework. *AIChE Journal* 2007, 53, 2832-2840.

207. Myers, A. L.; Prausnitz, J. M., Thermodynamics Of Mixed-Gas Adsorption. *AIChE Journal* 1965, 11, 121-127.

208. Reid, R. C.; Prausnitz, J. M.; Sherwood, T. K., *The Properties of Gases and Liquids*. 3rd ed.; McGraw-Hill: New York, 1977.

209. Jensen, C. R. C.; Seaton, N. A., An isotherm equation for adsorption to high pressures in microporous adsorbents. *Langmuir* 1996, 12, 2866-2867.

210. Valenzuela, D. P.; Myers, A. L.; Talu, O.; Zwiebel, I., Adsorption Of Gas-Mixtures - Effect Of Energetic Heterogeneity. *AIChE Journal* 1988, 34, 397-402.

211. Monson, P. A., On the Molecular-Basis of Adsorbed Solution Behavior. *Chemical Engineering Science* 1987, 42, 505-513.

212. Fairen-Jimenez, D.; Seaton, N. A.; Duren, T., Unusual Adsorption Behavior on Metal-Organic Frameworks. *Langmuir* 2010, 26, 14694-14699.

213. Sircar, S., Influence Of Adsorbate Size And Adsorbent Heterogeneity On IAST. *AIChE Journal* 1995, 41, 1135-1145.

214. Fischer, M.; Gomes, J. R. B.; Froba, M.; Jorge, M., Modeling Adsorption in Metal-Organic Frameworks with Open Metal Sites: Propane/Propylene Separations. *Langmuir* 2012, 28, 8537-8549.

215. Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M., Uff, A Full Periodic-Table Force-Field For Molecular Mechanics And Molecular-Dynamics Simulations. *J. Am. Chem. Soc.* 1992, 114, 10024-10035.

216. Li, H.; Eddaoudi, M.; O'Keeffe, M.; Yaghi, O. M., Design and synthesis of an exceptionally stable and highly porous metal-organic framework. *Nature* 1999, 402, 276-279.

217. Chui, S. S.-Y.; Lo, S. M.-F.; Charmant, J. P. H.; Orpen, A. G.; Williams, I. D., A chemically functionalizable nanoporous material [Cu-3(TMA)2(H2O)3]n. *Science* 1999, 283, 1148-1150.

218. Barthelet, K.; Marrot, J.; Ferey, G.; Riou, D., V^{III}(OH){O2C-C6H4-CO2}.(HO2C-C6H4-CO2H)x(DMF)y(H2O)z (or MIL-68), a new vanadocarboxylate with a large pore hybrid topology: reticular synthesis with infinite inorganic building blocks? *Chemical Communications* 2004, 520-521.

219. Barthelet, K.; Marrot, J.; Riou, D.; Ferey, G., A breathing hybrid organic-inorganic solid with very large pores and high magnetic characteristics. *Angewandte Chemie-International Edition* 2001, 41, 281-+.

220. Rosi, N. L.; Eddaoudi, M.; Kim, J.; O'Keeffe, M.; Yaghi, O. M., Infinite secondary building units and forbidden catenation in metal-organic frameworks. *Angewandte Chemie-International Edition* 2002, 41, 284-+.

221. Gelb, L. D.; Gubbins, K. E., Pore size distributions in porous glasses: A computer simulation study. *Langmuir* 1999, 15, 305-308.

222. Cessford, N. F.; Seaton, N. A.; Dueren, T., Evaluation of Ideal Adsorbed Solution Theory as a Tool for the Design of Metal-Organic Framework Materials. *Industrial & Engineering Chemistry Research* 2012, 51, 4911-4921.

223. Prosenjak, C. Experimental and theoretical adsorption studies in tunable organic-inorganic materials. The University of Edinburgh, Edinburgh, 2009.

224. Michels, A.; De Graaff, W.; Ten Seldam, C. A., Virial Coefficients of Hydrogen and Deuterium at Temperatures between -175-Degrees-C and +150-Degrees-C - Conclusions from the second Virial Coefficient with Regards to the Intermolecular Potential. *Physica* 1960, 26, 393-408.

225. Martin, M. G.; Siepmann, J. I., Transferable potentials for phase equilibria. 1. United-atom description of n-alkanes. *J. Phys. Chem. B* 1998, 102, 2569-2577.

226. Reid, R. C., Prausnitz, J.M., Sherwood, T.K., *The Properties of Gases and Liquids*. 3rd ed.; McGraw-Hill: New York, 1977.

227. Potoff, J. J.; Siepmann, J. I., Vapor-liquid equilibria of mixtures containing alkanes, carbon dioxide, and nitrogen. *AIChE Journal* 2001, 47, 1676-1682.

228. Dunne, J.; Myers, A. L., Adsorption of Gas-Mixtures in Micropores - Effect of Difference in Size of Adsorbate Molecules. *Chemical Engineering Science* 1994, 49, 2941-2951.

229. Jorge, M.; Lamia, N.; Rodrigues, A. E., Molecular simulation of propane/propylene separation on the metal-organic framework CuBTC. *Colloids and Surfaces a-Physicochemical and Engineering Aspects* 2010, 357, 27-34.

230. CLX compute cluster, School of Engineering, University of Edinburgh. http://www.see.ed.ac.uk/it/network/farm/clx/. (accessed October 2013).

231. Beowulf.org: The Beowulf cluster site. http://www.beowulf.org/. (accessed October 2013).

232. Condor compute pool, School of Engineering, University of Edinburgh. http://www.see.ed.ac.uk/it/network/farm/condor/. (accessed October 2013).

233. Edinburgh Compute and Data Facility web site, University of Edinburgh. http://www.ecdf.ed.ac.uk/. (accessed October 2013).